UTRECHT UNIVERSITY

Recent Advances in Phosphoproteomics: Principles and Applications

Name: Mariam Magdy Ahmed
Student ID: 3564231
Drug Innovation Master Program
Utrecht University
September 14th 2012 – October 25th 2012

Supervised by Dr. Dirk T.S. Rijkers, department of Medicinal Chemistry, Utrecht University

Reviewed by Dr. Shabaz Mohammed, group of Biomolecular Mass Spectrometry and Proteomics, Utrecht University

Summary

Protein phosphorylation is a reversible modification of cellular proteins that regulate protein function, cellular localization and formation of protein complexes. It is involved in almost all cellular processes and deregulation of protein phosphorylation has been implicated in the development of many diseases. Currently, mass spectrometry (MS) is the method of choice for studying cellular phosphorylation events, as the state-of-art MS techniques are able to analyze thousands of phosphopeptides present in complex biological samples in one experiment, in a rapid and highly sensitive fashion. However, phosphorylation analysis with MS is challenged by the substoichiometric levels of phosphorylated peptides, suppression of phosphopeptide ionization inside mass spectrometers and the loss of the phosphate group during peptide ion fragmentation in tandem mass spectrometry (MS/MS). With this respect, various developments in sample preparation strategies, instrumentation, quantitative methodologies and data analysis tools have been achieved to enhance MS-based phosphorylation analysis. Selective isolation and enrichment of phosphorylated proteins/peptides in analytical samples can be achieved prior to MS analysis by techniques such as, immobilized metal affinity chromatography (IMAC) using metal ions, metal oxide chromatography (MOAC) using metal oxides, other chromatographic methods such strong cation exchange (SCX) chromatography or the use of specific antibodies directed against specific phosphorylated residues in immunoprecipitation, while, combining different enrichment techniques are frequently employed to enhance the efficiency of enrichment and the subsequent phosphorylation analysis. Several MS approaches were developed to facilitate the analysis of phosphopeptides which can be difficult due the loss of the phosphate group. These include multi-stage MS with additional activation steps (e.g. MS^3), mass spectrometers with high-energy implementations and special scan modes and the employment of ion fragmentation techniques that avoid the phosphate group loss, such as electron-based dissociation techniques. In this literature review, the principles and applications of the state-of-art MS methodologies, current challenges and future directions in phosphoproteomics are described. In addition, some of the advances made in chemical biology tools for studying protein phosphorylation are discussed.

Table of Contents

1. Introduction **4**

1.1 Phosphoproteomics 4

1.2 Analysis of phosphorylation with mass spectrometry 5

1.3 Scope of this thesis 7

2. Sample preparation **8**

2.1 Enzymatic digestion 8

2.2 Enrichment techniques 9

2.2.1 Immobilized metal affinity chromatography (IMAC) 10

2.2.2 Metal oxide chromatography (MOAC) 13

2.2.3 IMAC versus MOAC and the development of SIMAC 15

2.2.4 Ion exchange chromatography 16

2.2.5 Hydrophilic interactions chromatography (HILIC) 18

2.2.6 Multi-dimensional chromatography 18

2.2.7 Calcium and barium precipitation 19

2.2.8 Immunoprecipitation 20

2.2.9 Chemical derivatization 21

2.2.10 Coupling of enrichment techniques 22

3. **Strategies of mass spectrometry in phosphoproteomics** 23

3.1 General principles of mass spectrometry 23

3.2 Ionization techniques 23

3.2.1 Electrospray ionization (ESI) 24

3.2.2 Matrix-assisted laser desorption ionization (MALDI) 24

3.3 Ionization of phosphopeptides 25

3.4 strategies of tandem mass spectrometry 25

3.4.1 Mass analyzers and hybrid mass spectrometers 26

3.4.1.1 Time-of-flight (TOF) instruments 26

3.4.1.2 Ion trap (IT) instruments 27

3.4.1.3 Fourier transform ion cyclotron resonance (FTICR) instruments 27

3.4.1.4 Orbitrap instruments 28

3.4.2 Phosphopeptide fragmentation 28

3.4.2.1 Collision-induced dissociation (CID) ... 29
3.4.2.1 Electron-based dissociation ... 31
3.5 Choice of instrumentation ... 32
4. Quantitative phosphoproteomics ... 34
4.1 Stable isotope labeling ... 34
4.1.1 Metabolic labeling ... 35
4.1.2 Chemical labeling ... 36
4.2 Label-free quantitation ... 35
5. Data analysis ... 38
5.1 Database search ... 38
5.2 Phosphosite localization ... 39
5.3 Data storage and mining ... 39
5.4 Sulfonation ... 40
6. Chemical biology tools to study protein phosphorylation ... 41
6.1 Peptide biosensors ... 41
6.1.1 Environmentally sensitive biosensors ... 42
6.1.2 Deep-quench biosensors ... 43
6.1.3 Self-reporting biosensors ... 43
6.1.4 Metal-chelation enhanced fluorescence biosensors ... 43
6.1.5 Genetically encoded protein sensors ... 44
6.2 Small molecule mimics of phosphoamino acids ... 45
6.3 Activity-based probes ... 46
7. Challenges and future directions ... 48

Chapter 1: Introduction

1.1 Phosphoproteomics

Phosphoproteomics is a research field in proteomic studies that is concerned with the broad scale studying of cellular protein phosphorylation events (1). Protein phosphorylation is the most extensively studied post translational protein modification (2) and strategies for studying this protein modification are in continuous developments, since it is a major player in many processes influencing cellular fate and function (3-5). Proteins of eukaryotic cells are mainly phosphorylated on serine, threonine or tyrosine amino acid residues by the enzymatic activities of protein kinases that catalyze the transfer of γ-phosphate group from adenosine 5'-triphosphate (ATP) to side chains of these specific amino acids in proteins (6). Protein phosphorylation is also regulated by another class of enzymes, phosphatases, which catalyze the removal of the phosphate group from phosphorylated proteins (7). Protein phosphorylation is a regulatory modification of protein function that is essentially involved in cellular signal transduction pathways that can mediate enzymatic activation or deactivation, recognition by other interacting proteins, formation of protein complexes or alteration of protein cellular localization and therefore plays a pivotal role in cellular processes such as cell cycle control, differentiation, proliferation and metabolism (2, 6, 8). Disrupted signaling pathways which can result from mutations in specific kinases or phosphatases are implicated in the pathology of several human diseases including cancer, autoinflammatory diseases, diabetes and neurodegenerative diseases (2, 8). If phosphorylation is a consequence of enzymatic activities of kinases and phosphatases, therefore understanding protein phosphorylation events in disrupted cellular signaling pathways of such diseases is essential to understand the underlying causes of the disease, the identification of potential drug targets and disease biomarkers that enable targeted treatment of patients and the interrogation of drug resistance mechanisms (8). Experimental techniques that are developed to quantify cell signaling, can measure and quantify protein phosphorylation as a direct assessment of signaling activity (3). Established techniques of measuring protein phosphorylation include radioactive ATP assay in which activity of kinases can be assessed by measuring the number of radiolabeled (^{32}P or ^{33}P) phosphates transferred from ATP carrying the radiolabel to the protein substrate (9, 10), the

use of specific antibodies directed against specific phosphorylated residues followed by immunoprecipitations and western blotting for identification (10) and several biochemical techniques that contributed significantly to the understanding of cell signaling (11). However, many of these techniques suffer from difficulty in automation, low throughput, some carry safety concerns such as radiolabeling and importantly limited by the complexity and dynamicity of phosphorylation events (4, 5, 9). The complex nature of phosphorylation events arise from the possibility of phosphorylating multiple sites on a single protein (4), with the presence of around 700,000 potential phosphorylated sites for any given kinase (6), where there are more than 500 kinases encoded from the human genome (12). Moreover, it is predicted that more than 30% of cellular proteins are phosphorylated at any given time (8). Investigating phosphorylation events in signaling pathways is challenging because of the aforementioned reasons, in addition to the branched networking of protein-protein communications that are tightly regulated by time and place (8). Currently, mass spectrometry (MS)-based strategies are the methods of choice for either characterizing single phosphorylated proteins or in performing global phosphoproteome analysis of analytical samples obtained from cells or tissues, for their speedy and highly sensitive analytical operations, with the ability to analyze thousands of phosphopeptides in a single experiment (3, 13). Moreover, MS is a powerful analytical technique for studying signaling networks, as it is able to quantify activities of all kinases present in a given proteome, and the identification and quantitation of phosphorylation sites with sufficient sensitivity and robustness (5, 8).

1.2 Analysis of phosphorylation with mass spectrometry

Detection and quantitation of protein phosphorylation with MS is analytically challenging because of several reasons (3, 13). For instance, the reversibility nature of protein phosphorylation makes it a transient modification that can be difficult to apprehend during analysis (13). Furthermore, when performing tandem mass spectrometry (MS/MS) experiments, where peptide ions are fragmented inside the mass spectrometer into product ions to gain structural information (13), ions are selected for fragmentation based on their intensities, which creates bias toward the analysis of the more abundant (i.e. higher intensity) peptides (3). This is of particular interest when analyzing phosphorylated peptides which are less abundant than their non-phosphorylated counterparts owing to their

substoichiometric levels (3). Substoichiometric levels of phosphorylated peptides mean that for a given peptide population the phosphorylated fraction is less than the non-phosphorylated major fraction (13). Moreover, within this small fraction of phosphorylated peptides different phosphorylation sites can exist, giving rise to different phosphorylated peptide isoforms (13). Furthermore, the ionization of phosphopeptides can be suppressed by their non-phosphorylated counterparts (14). Attempts to separate phosphopeptides from non-phosphorylated peptides before MS analysis were made by coupling a nano-flow liquid chromatography (nano-LC) system to mass spectrometers to reduce ion suppression (5, 14). However, the more hydrophilic multiply phosphorylated peptides may not bind efficiently to the hydrophobic stationary phase of the LC column, resulting in reduced recovery of phosphopeptides (14). Finally, the labile phosphate group can be lost during ion fragmentation in MS/MS experiments, making the generated phosphopeptide spectra to be more difficult to interpret than in the case of non-phosphorylated peptides (13-15). With this respect, various developments in sample preparation strategies, instrumentation, quantitation methodologies and data analysis tools have been achieved to enhance MS-based phosphopeptide detection, quantitation and identification of phosphorylation sites (13, 15, 16). For guidance through the coming chapters, Figure 1 shows an overview of the workflow performed in a typical phosphoproteomic study (17).

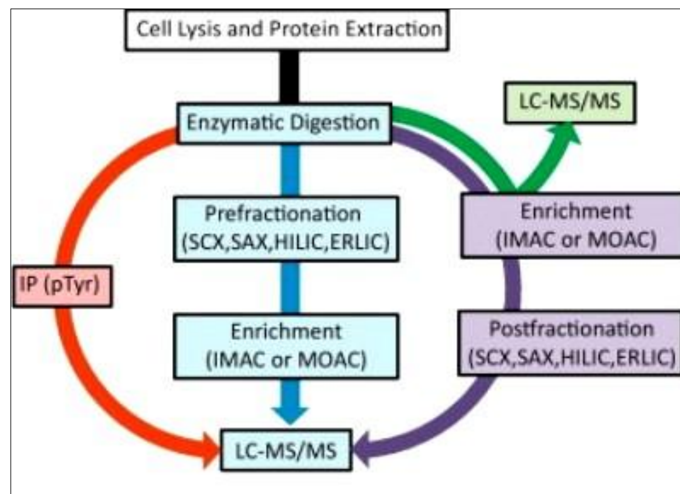

Figure 1 (17)**. An overview of the workflow performed in a typical phosphoproteomic study.** Cellular proteins are extracted by cell lysis and digested into a set of peptides using appropriate sequence-specific enzyme(s). Generated peptides are isolated and enriched in analytical samples using various enrichment techniques. Finally phosphopeptide-enriched samples are analyzed with mass spectrometry.

1.3 Scope of this thesis

This thesis describes recent MS approaches to facilitate and enhance the analysis of phosphorylation with MS. The main principles, some of the recent applications and current limitations of each approach are discussed. Because phosphoproteomics is a rapidly progressing research area with many reports published each year, the writing process was guided by these recent reviews (3-5, 13, 15-17). Advances in sample preparation strategies for phosphoproteomics including sample preservation, enzymatic digestion of specimen proteins, phosphopeptide separation and enrichment techniques are described in chapter 2, with examples of recent studies that employed such techniques. An important illustration provided with enrichment techniques, is the importance of using coupled separation and/or enrichment methods for complementary coverage of the phosphoproteome. Significant advances in mass spectrometry instrumentation and recent strategies in tandem mass spectrometry for phosphoproteomics are described in chapter 3. In addition, basic principles of mass spectrometry are provided for illustrative purposes. Quantitative phosphoproteomics including stable isotope labeling quantitation and label-free quantitation are described in chapter 4, illustrating the importance to differentiate between true changes that occur in the phosphorylation state of a cell after a specific stimulus, from those that can occur as a result of changes in protein abundance, which can be influenced by the experimental conditions. Data analysis tools available for interpretation of phosphopeptide spectra with the localization of phosphorylation sites are described in chapter 5, illustrating the need for parallel developments in software technologies with the developments made in sample preparation techniques and instrumentation that result in massive increase in phosphoproteomic data. Separate from mass spectrometry methodologies, some of the recent chemical biology tools that contributed to the current knowledge of protein phosphorylation are described in chapter 6. Finally, current challenges facing the field of phosphoproteomics and the future directions are discussed in chapter 7.

Chapter 2: Sample preparation

Sample preparation is very important to overcome the limitations encountered in MS-based phosphopeptide detection and analysis (3, 4, 18). For instance, the dynamic nature of phosphorylation requires sample preservation to protect cellular phosphorylation states, which can be achieved by adding acids (e.g. TFA) before protein harvesting to cease all enzymatic activities (18) and phosphatase inhibitors during cell lysis (19). Moreover, rapid denaturing of cellular proteins helps preserving analytical samples (19, 20). The low stoichiometry of phosphopeptides and their ionization suppression by the non-phosphorylated peptides can be overcome by selective isolation of phosphopeptides and their enrichment in analytical samples (3). The various techniques for phosphopeptide isolation and enrichment are discussed in this chapter.

2.1 Enzymatic digestion

For sequencing and subsequent identification of proteins by mass spectrometry it is essential to digest the proteins of interest into a set of peptides using sequence-specific proteases before introduction into the mass spectrometer (21). This has to do with the sequencing efficiency of mass spectrometers, which can sequence small peptides (up to 20 residues) more efficiently than intact proteins (21). Moreover, the solubility properties and ionization efficiencies of peptides are better than that of intact proteins (21). Enzymes that are chosen for proteomic experiments should exhibit high cleavage efficiency and specificity with the ability to generate peptides that are suitable for MS analysis (22). Trypsin is the most employed protease for protein digestion (3) and cleaves proteins at the C-terminal sides of lysine and arginine residues with high cleavage specificity generating a set of peptides that are suitably handled in most peptide fractionation techniques and subsequent analysis by tandem mass spectrometry (22). However, certain features of phosphorylated proteins can reduce its cleavage efficiency (3). For example, the possible interactions of the phosphate group with side chains of lysines or arginines and the repetitive occurrence of basic residues near phosphorylation sites can compromise the efficiency of trypsinization (3). Another factor that can restrict trypsin cleavage is the presence of proline residues near the cleavage sites (22). In phosphoproteomics, some alternative enzymes have been

introduced to provide data that is complementary to that generated by trypsin leading to the identification of more phosphopeptides such as the complementary use of the protease Lys-N (1). Lys-N is a recently introduced metalloendopeptidase enzyme (23) that cleaves at the N-terminal side of lysine residues and can efficiently function under high denaturing conditions (24). The complementary use of Lys-N with trypsin, generated populations of phosphopeptides with marginal overlap, where lys-N was able to identify 1759 unique phosphopeptides out of a total of 5036 identified phosphopeptides (1). Another protease, Lys-C, which cleaves specifically at the carboxylic side of lysine residues (22) was suggested to be useful for generating peptides that are characterized with electron transfer dissociation (ETD) (22). ETD is an ion fragmentation technique employed in tandem mass spectrometry that will be described in more details in chapter 3. The compatibility of the selected protease with the method of separation or enrichment is important, as for example tryspin was found to generate peptides that are not suitable with hydrophilic interaction chromatography (HILIC) because internal lysine and arginine residues are beneficial for peptide separation using this technique (25).

2.2 Enrichment techniques

By exploiting the properties of the phosphate group, various enrichment techniques can be employed (1). For example, the negative charge and polarity of the phosphate moiety can impart selectivity for separating phosphopeptides from other non-phosphorylated peptides which led to the employment of techniques such as strong cation exchange (SCX) chromatography, strong anion exchange (SAX) chromatography and hydrophilic interaction chromatography (HILIC) (26). Furthermore, the ability of the phosphate group to coordinate with metals and metal oxides has led to the development of techniques such as immobilized metal affinity chromatography (IMAC) (27) and metal oxide chromatography (MOAC) (28) respectively. In addition, calcium (29) and barium (30) precipitation of phosphopeptides provided alternative means of enrichment. The use of specific antibodies directed against phosphorylated amino acid residues can be used as an enrichment technique at the phosphoprotein level or at the phosphopeptide level (3). Moreover, chemical tagging of phosphate groups is yet another strategy of enrichment (16). A summary of all enrichment techniques of phosphorylated proteins/peptides are shown in Figure 2 (3).

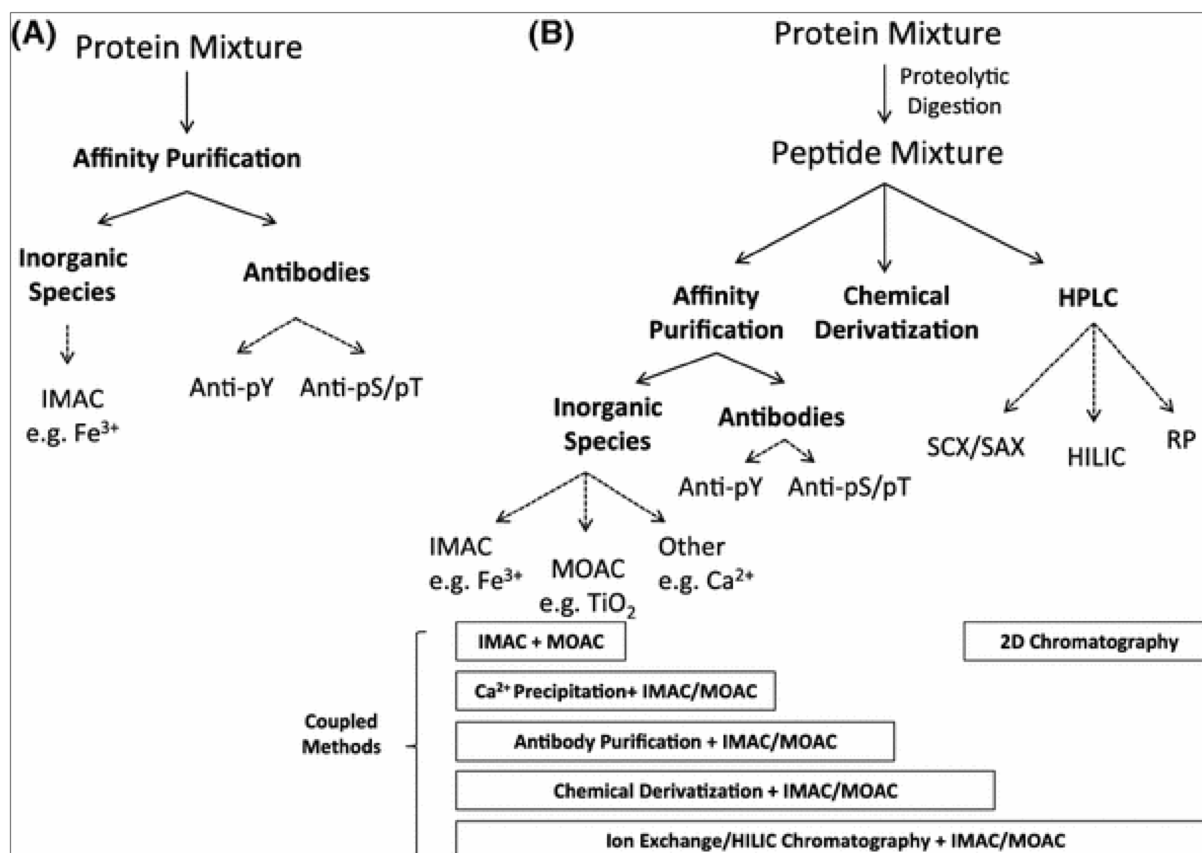

Figure 2 (3). Summary of enrichment strategies commonly used for phosphoproteins (a) and phosphopeptides (b) (3).

2.2.1 Immobilized metal affinity chromatography (IMAC)

The concept of using IMAC as a phosphopeptide enrichment strategy makes use of the ability of phosphate groups to interact with transition metal ions through the formation of coordinate bonds between oxygen atoms (electron donors) of the phosphate group and metal ions (electron acceptors) (3) (see Figure 3) (4). In an experimental context metal ions are immobilized onto porous column packing material to which phosphopeptide-containing complex samples are applied, and selective capturing of phosphopeptides mediates their isolation from the protein mixture (5). IMAC was first introduced for phosphopeptide enrichment in 1986 using trivalent iron (Fe^{3+}) ions (27), and many several contributions have followed using different metal ions including gallium (Ga^{2+}) (31, 32), Zinc (Zn^{2+}) (33) and zirconium (Zr^{4+}) (34); while Fe^{3+} remains the most commonly used for IMAC-based phosphopeptide enrichment for its higher affinity to phosphate groups compared to carboxyl and phenolic functional groups, imparting selective advantage (3).

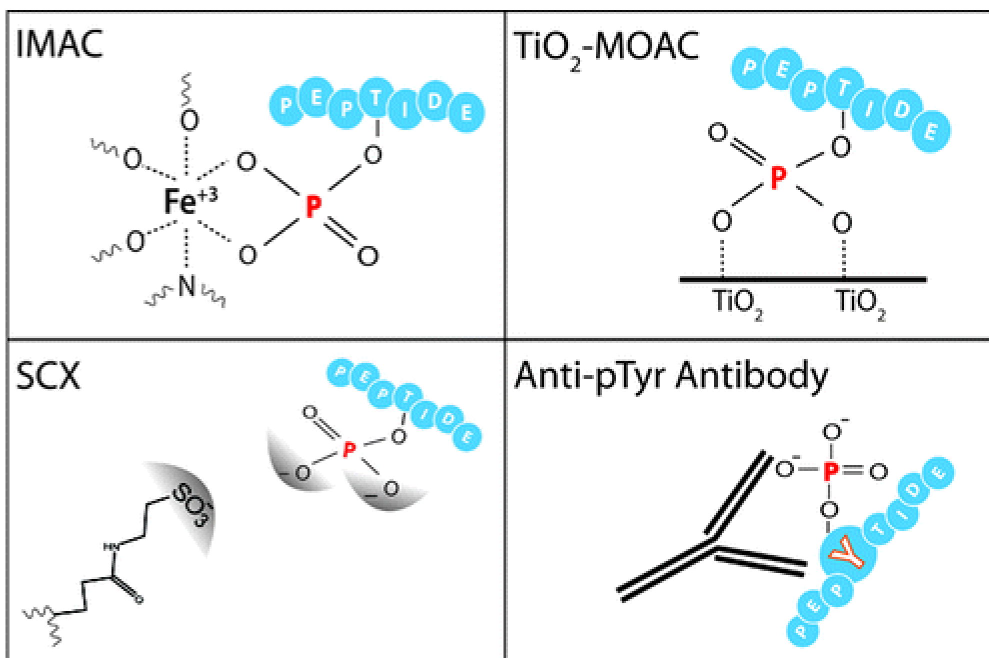

Figure 3 (4). Selective isolation of phosphopeptides can be mediated by the interaction of the phosphate group with Fe^{3+} ions in immobilized metal affinity chromatography (IMAC), TiO_2 in metal oxide chromatography (MOAC), sulfonic acid derivatives in strong cation exchange (SCX) chromatography or with specific antibodies directed against specific phosphorylated residues, such as anti-phosphorylated tyrosine antibodies (anti-pTyr).

IMAC (Fe^{3+}) has been extensively used for phosphopeptide enrichment in distinct applications including the determination of phosphorylated sites present in cystic fibrosis transmembrane conductance regulator protein whose activity is regulated by protein phosphorylation (35), *in vitro* investigation of the possible phosphorylation modification of the plant enzyme sucrose synthase (36), purification of chemically-synthesized phosphopeptides (37) and the characterization of digested phosphoproteins derived from *Saccharomyces cerevisiae* cell extracts (38). Although various applications of IMAC for phosphopeptide enrichment have been implemented as described above, nevertheless, it suffers from certain limitations, mainly the aspecific binding of other functional groups such as carboxyl groups of acidic amino acids, comprising glutamic acid, aspartic acid, histidine and cysteine resulting in the co-recovery of acidic peptides (3, 39). The co-recovery of acidic peptides is more significant in wide proteomic studies, where the specificity to phosphoproteins is suboptimal (60-70 %) (39). However, the aspecific binding of acidic peptides to IMAC material can be reduced by several strategies including methyl esterification of peptide carboxyl groups before IMAC enrichment to shield their interaction with metal ions, which resulted in the identification of higher number of phosphopeptides

and phosphorylation sites from whole lysates (38, 40-42). However, it has been reported that chemical modifications of peptides resulted in severe sample loss because of the multi-step procedures (39, 43). The selective recovery of phosphopeptides using IMAC can be optimized by controlling the experimental conditions, such as pH and ionic strength of experimental buffers and solvents (3). Modulation of pH is one of the critical parameters influencing the specificity of IMAC to phosphopeptides by dictating the degree of protonation and deprotonation of functional groups including phosphates, and therefore influencing the interaction with metal ions (39). For example, at high pH values (above 3.5) phosphate groups become completely deprotonated and exhibit enhanced binding to metal ions (39). However, this increased binding of phosphopeptides is at the expense of reduced selectivity of IMAC material, since carboxylic groups are concomitantly deprotonated and binds to IMAC resulting in the co-recovery of acidic peptides (39). The careful choice of acids used in washing buffers can greatly influence the aspecific binding to IMAC material, as it was demonstrated that high concentrations of acetic acid significantly reduced the aspecific binding with the recovery of higher number of phosphopeptides compared to formic acid which can competitively coordinate with Fe^{3+} (39). Therefore, it was recommended that acids like acetic acid and trifluoroacetic (TFA) acid are better included in the washing buffers, while formic acid is better included in elution buffers (39). Furthermore, it was shown that using TFA in 0.1% concentration in the loading buffer resulted in reduced binding of mono-phosphorylated peptides to the IMAC material, with the retention of multi-phosphorylated peptides on the column (44). Moreover, some salts such as ammonium salts commonly used in the trypsinization buffer, can negatively influence the binding of phosphopeptides to IMAC material through the formation of "ion pairs" with phosphopeptides (39). Replacing buffers containing ammonium salts with triethyl ammonium bicarbonate buffer enhanced the binding of phosphopeptides by eliminating the possible ionic interactions with phosphopeptides (39). Furthermore, the degree of ionization of carboxylic and phosphate groups were investigated at different concentrations of acetonitrile in water upon application to IMAC (Fe^{3+}) and it was demonstrated that at higher concentrations of acetonitrile, the ionization of carboxylic groups was reduced while that of phosphate groups was not changed resulting in increased specificity of IMAC materials toward phosphopeptides (45). IMAC (Fe^{3+}) has shown to preferentially enrich multi-phosphorylated peptides than singly phosphorylated peptides (44, 45) however, the elution

conditions can greatly influence the type of enriched phosphopeptides (3). For example, acidic elution conditions using increasing concentrations of TFA recovered more mono-phosphorylated peptides compared to multi-phosphorylated peptides, whilst basic conditions using ammonium hydroxide favored the recovery of multi-phosphorylated peptides (44). As previously mentioned, phosphoproteins can be enriched at either the protein level or at the peptide level, while it was reported that IMAC was used to enrich phosphorylated proteins at both levels in the same experiment in a method named "double IMAC" resulting in the identification of distinct phosphopeptides and phosphorylation sites in the epidermal growth factor pathway (46). Repetitive enrichment cycles at the peptide level have been performed in a single experiment resulting in efficient recovery of phosphopeptides (45).

2.2.2 Metal oxide chromatography (MOAC)

Metal oxide chromatography (MOAC) was introduced as an alternative to IMAC to increase phosphopeptides selectively (3). Phosphopeptide enrichment via MOAC is based on the ability of phosphate groups to interact with metal oxides (see Figure 3) (4), and the technique was first introduced with the use of a TiO_2 (titanium dioxide) column coupled to reverse-phase liquid (RP) chromatography (28). Test peptides were applied to the TiO_2 column under acidic conditions (pH 2.9) to permit binding of phosphopeptides and were recovered with alkaline conditions (pH 9.0) (28). After recovery from the TiO_2 column, phosphopeptides were concentrated on the RP-LC column and were analyzed with MS/MS (28). The authors demonstrated the ability of their novel enrichment method to separate phosphopeptides from non-phosphorylated peptides at 90% enrichment yield however, they also observed nonspecific binding of acidic peptides and they suggested methyl esterification to reduce the nonspecific binding (28). Since then, TiO_2-based phosphopeptide enrichment gained great popularity in proteomic applications (3) and was employed in many studies, including the investigation of cellular responses toward aberrant signaling resulting from mutated receptor tyrosine kinase Flt3 in a wide-scale phosphoproteomic study (47). In this study, phosphopeptides of one sample were enriched using multiple cycles of TiO_2-based enrichments, while other samples were enriched using strong cation exchange chromatography prior to TiO_2-based enrichment, and the authors were able to identify more than 14700 phosphorylation sites (47). In another study, TiO_2-based

phosphopeptide enrichment was investigated under variable experimental conditions of peptide loading and elution (48). Selective enrichment of phosphopeptides with low recovery of non-phosphorylated peptides was achieved by loading and washing phosphopeptide-containing mixtures under organic acidic conditions (48). Increasing the organic acid content in binding buffers was positively correlated with phosphopeptide binding with optimal selective isolation at 1% TFA when compared to other organic acids as formic acid and acetic acid (48). High phosphopeptide recovery with increased selectivity was achieved by elution under basic conditions using ammonium hydroxide or ammonium phosphate (48). Furthermore, the authors didn't recommend the use of glycolic acid in loading buffers as a non-phosphopeptide excluder, which was previously reported to enhance phosphopeptide selectivity to TiO_2 columns (49), as they found it to reduce the phosphopeptide selectivity rather than its enhancement (48). In addition, phosphopeptides generated by trypsin digestion of the model α- and β-casein phosphorylated proteins were enriched with different IMAC and MOAC materials: MOAC (TiO_2), MOAC (ZrO_2), IMAC (Fe^{3+}) and IMAC (Ga^{2+}) where peptides were loaded into all columns with 1% TFA and 30% acetonitrile, eluted with 0.4 M ammonium hydroxide in 30% acetonitrile and subsequently analyzed with mass spectrometry (48). In terms of selectivity, MOAC (TiO_2) was found to be more selective to phosphopeptides than MOAC (ZrO_2) and both IMAC materials which resulted in the co-elution of non-phosphorylated peptides (48). Furthermore, MOAC (TiO_2) was found to be more efficient in recovering multi-phosphorylated peptides compared to MOAC (ZrO_2) while both were similarly efficient in recovering singly phosphorylated peptides (48). However, more efficient recovery of multi-phosphorylated peptides was achieved with both Fe^{3+} and Ga^{2+} compared to TiO_2, which could be attributed to the higher binding affinity of multi-phosphorylated peptides to TiO_2 (3, 44, 48). Furthermore, strong elution conditions (e.g. high pH and high ionic strength) can enhance the recovery of multi-phosphorylated peptides from TiO_2 columns (3). While comparing IMAC (Fe^{3+}) and IMAC (Ga^{2+}), Fe^{3+} showed more bias toward the recovery of multi-phosphorylated peptides compared to Ga^{2+} (48).

2.2.3 IMAC versus MOAC and the development of SIMAC

As mentioned above, phosphopeptide selectivity of TiO_2 is higher than that of Fe^{3+} and Ga^{2+} (48). Jensen and coworkers have compared IMAC and TiO_2 performance in terms of phosphopeptide sensitivity and selectivity under the influence of salts and detergents routinely used in cell biology experiments (49). They demonstrated that IMAC performance was greatly influenced by such components and that desalting of samples is essential prior to IMAC enrichment (45, 49). In contrast, TiO_2 performance was less sensitive to the presence of salts and detergents, especially with the inclusion of non-phosphorylated peptides excluders such as 2,5-dihydroxybenzoic acid (DHB), phthalic acid and glycolic acid (49). However, DHB and phthalic acid were found to interfere with MS-based phosphopeptide analysis due to their retainment in the LC column and inlet of the mass spectrometer, which can cause damaging to the electrospray system (48, 50). Moreover, as previously mentioned glycolic acid reduces the phosphopeptide sensitivity of TiO_2 (48). For the aforementioned reasons, pre-purification steps are recommended prior to both enrichment strategies (3). TiO_2 chromatography was demonstrated to prejudicially recover singly phosphorylated peptides than multi-phosphorylated peptides, owing to the tighter binding of the multi-phosphorylated peptides to TiO_2 columns making them more difficult to elute (44, 48), while IMAC was found more efficient in recovering multi-phosphorylated peptides (48). Acidic peptides were found to bind both IMAC (39) and TiO_2 (28) materials however, TiO_2 exhibit higher phosphopeptide selectivity compared to IMAC (48). The detection of multi-phosphorylated peptides with MS is more difficult than mono-phosphorylated peptide, because their ionization can be suppressed by mono-phosphorylated and non-phosphorylated peptides (16, 44). A method referred to as sequential elution from IMAC (SIMAC) has been developed by combining IMAC and TiO_2 for the enrichment of both mono- and multi-phosphorylated peptides (44). SIMAC is performed in multi-steps starting with IMAC where two elution steps are performed, firstly with acidic elution conditions (pH 1) to recover mono-phosphorylated peptides and secondly with basic elution conditions (pH 11.3) to recover multi-phosphorylated peptides. Phosphopeptide selectivity of only the acidic eluent containing mono-phosphorylated peptides is enhanced by a second enrichment step with TiO_2 (44). This new SIMAC methodology was performed on human mesenchymal stem cell extracts and phosphopeptide enrichment efficiency was

compared to that of TiO_2 chromatography. The authors demonstrated that SIMAC was able to identify more than twice the number of phosphorylation sites obtained with TiO_2 (44).

2.2.4 Ion exchange chromatography

Ion exchange chromatography is a technique that is used for peptide separation, which depends on the charge-charge interactions between peptide analytes and a charged stationary phase for separation (26). Peptides are eluted by a salt gradient that competitively displaces the peptides from the stationary phase, so that the higher the charge of the peptide the stronger the interaction with the stationary phase and the later it is eluted (26). In phosphoproteomics, ion exchange chromatography was employed for isolating phosphopeptides from their non-phosphorylated counterparts by depending on the net negative charge of the phosphate group (3, 26). Two techniques in ion exchange chromatography were proposed for phosphopeptides enrichment including strong cation exchange (SCX) and strong anion exchange (SAX) (3). In SCX the stationary phase has negatively charged strongly acidic functional groups, most commonly sulfonic acid derivatives that can bind to positively charged peptide analytes (26) (see Figure 3) (4). The low pKa of the sulfonic acid functional group allows it to maintain its negative charge at low pH and therefore, stationary phase with sulfonic acid derivatives can be utilized over a wide pH range (26). An example of using SCX for enrichment of phosphopeptides is the work of Ballif et al (51). In this study, proteins extracted from developing brain tissues of a mouse embryo were trypsinized and separated on a SCX column at pH 3. At this pH value, the tryptic digests would contain: mostly protonated carboxylic groups that don't contribute to the net charge, positively charged terminal amino groups and positively charged basic amino acid residues (lysine, histidine and arginine) (51). The presence of a negatively charged phosphate group ($H_2PO_4^-$) will therefore decrease the net charge of the peptide by a difference of one compared to the non-phosphorylated counterpart (see Figure 4) and therefore the binding strength to the negatively charged stationary phase is less and the phosphopeptide would elute faster (51).

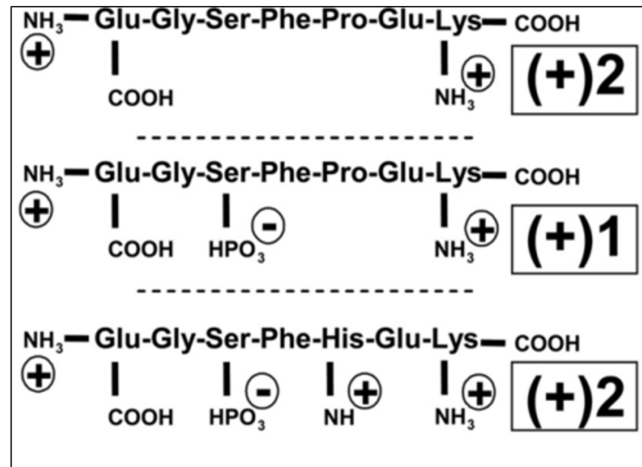

Figure 4 (51). **Schematic representation of tryptic peptide digests separated by SCX at pH 3**. The presence of the phosphate group reduces the net charge of the phosphopeptide by 1 compared to the non-phosphorylated peptide counterpart imparting selective separation of phosphopeptides from non-phosphorylated peptides based on net charge.

The authors were able to identify over 500 phosphorylation sites in the developing mouse brain (51). However, it was later demonstrated that acidic peptides were co-eluted with phosphopeptides and that hydrophobic phosphopeptides could be retained in the SCX column (52). Moreover, mono-phosphorylated peptides having more than one basic residue have the same net charge as non-phosphorylated peptides allowing their co-elution (26). Combining IMAC or TiO_2 with SCX is common to enhance phosphopeptide enrichment (53, 54). Furthermore, it was demonstrated that the use of the digestive protease Lys-N allowed the generation of peptides in which N-acetylated peptides, non-modified peptides and phosphopeptides have different net charges that could be separated on a SCX column (1, 23).

In strong anion exchange chromatography, the stationary phase has positively charged strongly basic functional groups (e.g. trimethylamino) that can be employed for phosphopeptide separation from non-phosphorylated peptides by taking advantage of the acidity of the phosphate group (26). The pH of the working buffers can influence the ionization form and the negative charge of the phosphate group, where phosphopeptides can acquire a net negative charge that mediate their separation from non-phosphorylated peptides on a SAX column (3, 26). In a study, phosphopeptides obtained from tryptic digests of the model phosphoproteins proteins α-casein and β-casein, were found to strongly bind SAX column (55). Furthermore, the number of identified phosphopeptides using SAX column was almost the same to that identified using IMAC (Fe^{3+}). However, SAX showed less bias for

multi-phosphorylated peptides and the complementary use of both techniques was suggested by the authors (55).

2.2.5 Hydrophilic interactions chromatography (HILIC)

Hydrophilic interactions chromatography (HILIC) can be used for phosphopeptide enrichment by making use of the hydrophilicity of the phosphate group (3). The mechanism of HILIC-based separation is not completely understood however, it was explained that by using a hydrophilic stationary phase and a low polar mobile phase (5-20% water in acetonitrile), a still aqueous layer is formed around the stationary phase where the analytes partition between the mobile phase and this aqueous layer, while elution is performed by increasing the concentration of water in the mobile phase (26). However, using HILIC as a solo technique for phosphopeptide enrichment is not efficient, since other non-phosphorylated peptides were found to co-elute however, it can be applied successfully when combined with other enrichment techniques such as IMAC or TiO_2 (3, 25).

2.2.6 Multi-dimensional chromatography

The coupling of high-pressure liquid chromatography (HPLC) with mass spectrometers (LC-MS) has become a standard in many proteomic studies to reduce sample complexity before introduction to mass spectrometers, and thus enhance the sensitivity of mass spectrometric measurements (26). Efficient separation of peptides can influence the sensitivity of MS measurements by the ability to separate the most abundant from the less abundant peptides and introduce them separately for sequencing (26). This is of particular interest with phosphopeptides which as previously mentioned, exist in substoichiometric levels compared to their non-phosphorylated counterparts and thus could be missed during MS measurements. In two-dimension LC configurations (using two separation techniques) the second dimension is usually reverse phase (RP) chromatography (26). In RP chromatography, peptides are separated according to the degree of hydrophobicity and hydrophilicity, with the use of a hydrophobic stationary phase and a polar mobile phase, and elution is performed by gradually increasing the organic solvent concentration in the mobile phase (26). RP chromatography is useful in the removal of salts from analytical samples especially when the first dimension technique requires salts for separation (26). Combining chromatographic techniques in multidimensional can provide orthogonality or

non-overlapping peptide separation (26). An example of using two-dimensional separation is the use of SAX as a first dimension and RP as the second in an online setup For LC-MS/MS analysis (56). Online setups allow the un-intervened transfer of analytes from one separation column to another (26). This setup was used for phosphoproteomic analysis of tryptic digests obtained from HeLa cells and the authors were able to identify 1497 unique phosphorylation sites (56). In another study, the authors aimed to separate and enrich basic phosphopeptides by employing two dimensional SCX chromatography, where in the first SCX dimension a pH of 3 was employed, whilst the pH in the second SCX was 1 (57). By changing the pH between the two SCX separations the charge of the phosphate group changes where at pH 3 it acquires a negative charge while at pH of 1 it becomes neutral. This change in the net charge will only take place in phosphopeptides, while the net charge of non-phosphorylated peptides remains unchanged imparting selective isolation of phosphopeptides (57). By using this multidimensional configuration the authors were able to identify over 10,000 basic phosphopeptides (57).

2.2.7 Calcium and barium precipitation

Calcium precipitation of phosphopeptides can be used as a complementary enrichment technique to other well established techniques (e.g. IMAC) for more efficient phosphopeptide enrichment (29). In this study, the investigators isolated phosphorylated peptides from tryptic digests of rice embryo proteins with calcium phosphate precipitation using a solution of disodium phosphate (Na_2HPO_4), ammonia solution ($NH_3.H_2O$) and calcium chloride chloride ($CaCl_2$) prior to IMAC (Fe^{3+}) enrichment (29). This combination increased phosphopeptide selectivity with the identification of 227 unique phosphorylation sites (29).

Another study provided barium (Ba^{2+})/acetone precipitation as a novel method for phosphopeptide enrichment (30). The authors first measured the binding properties of Ba^{2+} ions to phosphopeptides using a simple mixture of casein phosphopeptides. Then they applied Ba^{2+}/acetone based phosphopeptide precipitation using a more complex sample of HeLa cell nuclear extracts and the analysis was performed using Multidimensional Identification Technology (MudPIT) a strategy that was previously developed for large scale proteomic studies (58). The authors were able to identify 1037 phosphopeptides using 250 µg of HeLa cells nuclear extracts (30).

2.2.8 Immunoprecipitation

Specific antibodies directed against phosphorylated amino acid residues, can be used to immunoprecipitate phosphorylated proteins/peptides (3). Immunoprecipitated proteins can be separated by gel electrophoresis, stained with a protein gel stain, excised from the gel and analyzed with mass spectrometry (16). Because tyrosine phosphorylation is the least abundant in cells (59), it is difficult to enrich proteins with phosphorylated tyrosine residues (3). Antibodies directed against phosphorylated tyrosine (pY) (see Figure 3) (4) are therefore able to enrich phosphorylated tyrosine-containing peptides more efficiently than other methods (3), which is of particular interest in applications especially concerned with tyrosine phosphorylation (59). Practically, targeting pY with specific antibodies is more easy than targeting phosphorylated serine (pS) or phosphorylated threonine (pT) residues (3), as there more available anti-pY antibodies that have successfully immunoprecipitated pY-containing proteins (59), while antibodies directed against pS and pT are less selective (3). Furthermore, the aromaticity aspect of tyrosine residues imparts more modes of interactions with binding sites of antibodies facilitating the generation of Y-specific antibodies, in contrast to the aliphatic serine or threonine amino acids that provide less available modes of interactions with antibodies (3). However, the applicability of antibody-based enrichments is limited by the presence of protein motifs specifically recognized by antibodies and the commercial availability of specific antibodies (3, 60). In a study, the authors aimed to study the epidermal growth factor receptor signaling pathway by immunoprecipitating HeLa cellular proteins using several anti-pY antibodies with and without stimulation with EGF (61). Immunoprecipitated proteins were resolved by gel electrophoresis and were silver stained. MS/MS analysis of the excised bands identified 9 tyrosine-phosphorylated proteins involved in EGFR pathway, where two of them were not previously identified, including the guanosine nucleotide exchange factor (Vav-2) which was confirmed by this study to be involved in the EFGP signaling pathway and another pY-containing protein (61). However, their work was criticized by depending on the relative staining intensities of stimulated and non-stimulated protein bands for quantitation which was prone to bias and inaccuracy (5).

Several reports were published with the use of antibody-based phosphoprotein enrichment at either the peptide level or the protein level that are summarized in many reviews

including (3, 16). Combination of immunoprecipitation with IMAC or MOAC has also been reported in the literature (16).

2.2.9 Chemical derivatization

MS-based phosphopeptide analysis is challenged by the loss of the phosphate group during ion fragmentation using collision-induced dissociation (CID) and the poor ionization of phosphopeptides compared to their non-phosphorylated counterparts (13, 15). The poor ionization of phosphopeptides is observed in the positive ion mode MS, where the employed acidic conditions suppress the ionization of phosphopeptides because of the negatively charged phosphate that capture protons present in acidic buffers (14, 15, 62). Therefore, methods to remove or replace phosphate groups with more stable chemical tags before phosphopeptide enrichment can be employed to enhance MS analysis of phosphopeptides (3). In chemical derivatization, the phosphate group can be removed by β-elimination resulting in a reactive dehydroamino acid that can react with a chemical tag by Michael addition (3, 15). The includsion of affinity tags (e.g. biotin) with the chemical tags allow the selective isolation of phosphopeptides or proteins (3). In a study aimed at analyzing phosphoserine and phosphothreonine-containing peptides (63), phosphorylated serine and threonine of three model phosphoproteins (α-casein, β-casein and ovalbumin) were chemically derivatized into S-(2-mercaptoethyl)cysteinyl or β-methyl-S-(2-mercaptoethyl)cysteinyl residues by β-elimination performed under alkaline conditions (barium hydroxide solution), followed by Michael addition with 1,2-ethanedithiol (63). The chemically modified peptides were biotinylated, digested with trypsin, enriched using immobilized avidin and were analyzed by tandem mass spectrometry (63). All expected phosphorylation sites in the three model phosphoproteins were identified using this strategy (63). However, chemical derivatization suffer certain limitations including the amenability of β-elimination to side reactions, in addition to the slow reaction rate of phosphorylated threonine compared to phosphorylated serine, while phosphorylated tyrosine is left unreacted (64). Moreover, β-elimination is not entirely selective for phosphorylated serine and threonine residues, since glycosylated serine and threonine residues were also (65) found to be amenable to β-elimination (65) as well as hydroxyl groups of unmodified serine and threonine residues (66).

2.2.10 Coupling of enrichment techniques

Coupling different phosphopeptide enrichment techniques enhance the efficiency of enrichment, leading to the identification of more non-redundant phosphopeptides (3). Combining SCX as a pre-fractionation step with TiO_2 for further enrichment has been successfully used, including a study performed for selective enrichment of phosphopeptides from protein tryptic digests obtained from *Drosophila melanogaster S2* cell lysates (67). Coupling of SCX with TiO_2 in an automated online setup allowed the identification of 2152 phosphopeptides from 250 µg starting material (67). However, coupling of SAX with TiO_2 was suggested a better combination than SCX/TiO_2 providing more coverage of the phosphoproteome (68). In another study, SCX was coupled to SIMAC to provide enhanced phosphoproteome coverage of tryptc digests obtained from murine fibroblast cells (NIH-3T3 cell line), where SCX was used as a pre-fractionation step followed by SIMAC enrichment (69). Using this strategy, the authors were able to identify thousands of phosphopeptides from NIH-3T3 cell lysates (69).

Chapter 3: Strategies of mass spectrometry in phosphoproteomics

3.1 General principles of mass spectrometry for proteomics

Peptide mixtures obtained after enzymatic digestion of intact proteins are first fractionated by a nano-LC system that is usually based on reversed-phase peptide separation before introduction to the mass spectrometer to reduce sample complexity (5). The LC system can be connected to the mass spectrometer in an online setup to directly introduce the eluted peptides for mass spectrometric analysis (21). Typically mass spectrometers are based on three main components: an ion source, a mass analyzer and a detector (70). In an ion source polar non-volatile peptide species are converted into gas-phase intact ions that can be directed inside a mass spectrometer by electric or magnetic fields (70). The generated peptide ions are then introduced to the mass analyzer to be separated according to their mass to charge ratios (m/z) followed by their direction to an ion detector to record the m/z and intensity of each peptide ion in a full scan MS generating a mass spectrum showing the intensity of the precursor ion at each m/z ratio (4, 21, 70). The intensity of each peak indicates the abundance of the peptide species, such that the higher the peak intensity the higher the abundance of the peptide ion (4). Mass spectrometers can be used for straightforward measurements of peptide molecular weights in a single-stage MS mode or can be used for more complicated analysis in tandem mass spectrometry (MS/MS) to obtain structural information, such as primary amino acid sequences or type and site of protein modifications (70, 71). In MS/MS after the initial MS scan, ions are selected by descending abundances for fragmentation into product ions and analysis of the ion fragments provides the structural information of peptides (4, 71). Some instruments of mass spectrometry can perform repetitive number (n) of ion selection and fragmentation MS cycles in multistage MS/MS (MS^n) to obtain more informative sequences from peptides (13, 15).

3.2 Ionization techniques

The generation of intact non-fragmented peptide ions is achieved by “soft ionization” techniques (21) where the two most widely used are electrospray ionization (ESI) (72) and matrix-assisted laser desorption/ionization (MALDI) (73).

3.2.1 Electrospray ionization (ESI)

In ESI (Figure 5) (21), the liquid peptide analyte exit the liquid chromatography column to travel through a capillary that is placed at high electric potential relative to the inlet of the mass spectrometer causing the formation of charged droplets of the liquid analyte in a process known as "pneumatic nebulization" (21, 70, 74). Under the effect of the electric potential applied to the capillary when the liquid droplets reach the end of the capillary, they are sprayed into fine droplets from which liquid evaporates by hitting a heating device and/or through the formation of very small droplets by repetitive collisions (21, 70). Multiply charged bare peptide ions are subsequently formed by ion desorption from the droplet (21, 70). Reducing the flow rate of the liquid analyte from which electrosprayed ions are generated can enhance the efficiency of ESI which can be achieved by using small diameter capillaries and small diameter capillary orifices (70).

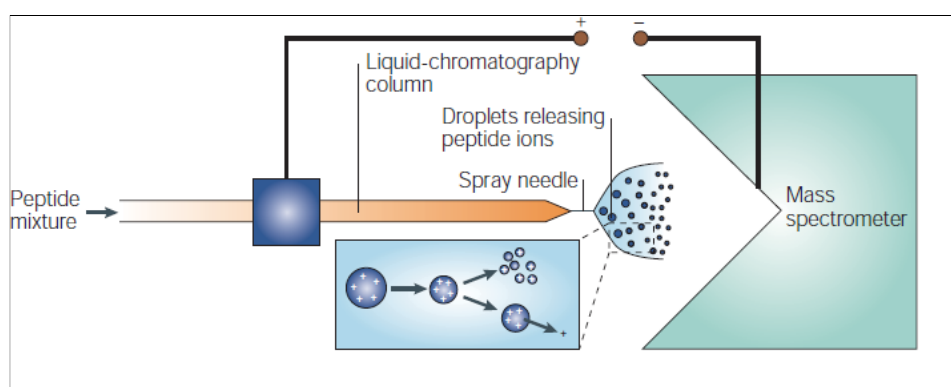

Figure 5 (21). Schematic representation of peptide ionization with electrospray ionization.

3.2.2 Matrix-assisted laser desorption ionization (MALDI)

In MALDI (Figure 6) (21), the ionization of the liquid peptide analyte is “assisted” by the presence of a matrix consisting of a low molecular weight aromatic acid that is able to absorb applied laser energy (21, 70). The technique is based on mixing the liquid analyte with the acidic matrix followed by promoting co-crystallization of the sample and matrix molecules (21, 70). Pulsing the crystals with a laser beam allow the matrix molecules to gain thermal energy resulting in their sublimation, taking with them the analyte molecules into the gas phase (21, 70). Desorbed ions and molecules in the gas phase undergo several ion-molecule collisions generating singly charged analyte ions that are directed into the mass spectrometer by electric potentials (21).

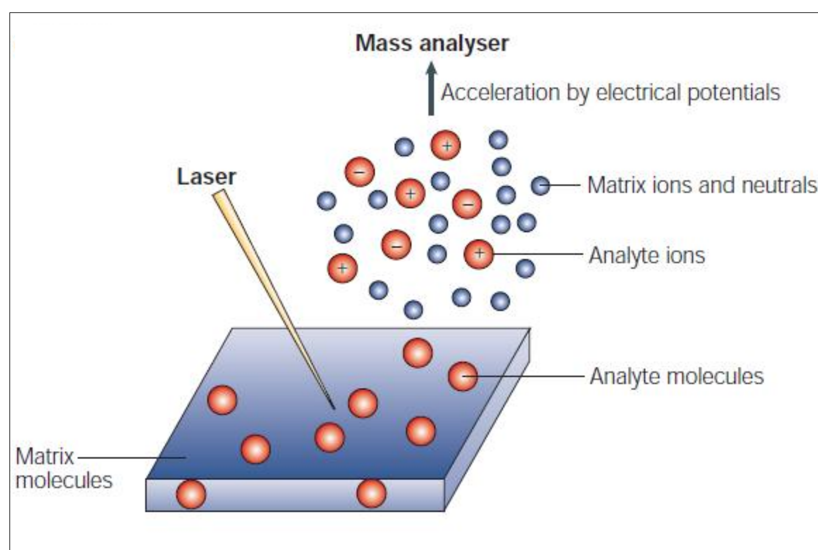

Figure 6 (21). Schematic representation of peptide ionization with matrix-assisted laser desorption ionization.

3.3 Ionization of phosphopeptides

It was proposed that phosphopeptides may exhibit less efficient ionization than non-phosphorylated peptides when using MALDI as the ionization technique (13). However, the ionization and subsequent detection of phosphopeptides in MS instruments coupled to MALDI can be enhanced by the use of matrix additives (13, 15) such as alkylphosphonic acids (75). When using ESI no specific inhibition of phosphopeptides ionization was demonstrated; however the lower abundance of phosphopeptides compared to non-phosphorylated peptides is what interferes with phosphopeptide MS detection (13).

3.4 Strategies of tandem mass spectrometry

Measuring the mass to charge ratios of intact ions is insufficient to identify a peptide, which can be achieved by sequencing, or to define the type and site of post translational modifications (PTMs) such as protein phosphorylation (70, 71). To gain such information, tandem mass spectrometry (MS/MS) strategies are performed (70, 71). MS-based phosphoproteomic studies aim to 1) specifically capture phosphorylated peptides from a rich pool of cellular peptides generated by enzymatic digestion and assign their identity from their amino acid sequences, 2) spot the specific phosphorylation sites, and 3) quantitatively measure changes in the phosphorylation state under certain biological manipulations specified by the interest of the investigator (13, 15). Because of the chemical nature of phosphate groups which impart difficulties in their detection and identification

inside a mass spectrometer (13), advances in MS/MS methodologies have evolved over the years to increase the efficiency of phosphopeptide characterization with high sensitivity, allowing the detection of phosphopeptides even at levels as low as the attomole level (76).

3.4.1 Mass analyzers and hybrid mass spectrometers

Different mass analyzers exist that differ in the way of measuring the m/z ratios, accuracy of mass measurements, resolution and operation speed including: time-of-flight (TOF), quadrupole (Q), ion trap (IT), orbitrap and fourier transfer ion cyclotron resonance (FTICR) (77). Performance of mass spectrometers can be evaluated in terms of accurate mass measurements and mass resolution which influence the capability of the mass spectrometer to identify charge states of ions and sort ions of the same charge (70). Different mass analyzers can be combined in "hybrid mass spectrometers" such as quadrupole time-of-flight (Q-Q-TOF) and tandem time-of-flight (TOF-TOF) to combine the qualities of mass analyzers which be used in performing specific analytical strategies (71, 77). The general principles of different mass analyzers and properties of hybrid mass spectrometers will be briefly discussed in this section as an introduction to the choice of instrumentation in phosphoproteomics which will be discussed later. Detailed descriptions of MS instrumentation can be found in these reviews (70, 71).

3.4.1.1 Time-of-flight (TOF) instruments

In TOF mass analyzers, the m/z ratios of the generated ions can be determined by measuring the flying time of an ion inside a specific tube with known length under vacuum (71). TOF analyzers can be an integral part of hybrid mass spectrometers to achieve high resolution and mass accuracy such as the Q-Q-TOF mass spectrometers which combine two quadrupole (Q) mass analyzers with a TOF mass analyzer (71). In MS/MS experiments Q-Q-TOF is efficient in quantitative proteomics and in identification of PTMs (71). Precursor ions are selected in the first quadrupole, fragmented into product ions via collision-induced dissociation (CID) in the second quadrupole and mass analysis of product ions is carried out by the TOF analyzer (71). MALDI is frequently coupled to TOF mass spectrometers for mass analysis (71, 77) and is also coupled with tandem (TOF/TOF) mass spectrometers (78) and Q-Q-TOF for tandem mass spectrometry analysis (71). In TOF/TOF instruments, ions are

selected in a TOF analyzer, fragmented in a collision cell and analyzed in a second TOF analyzer (70).

3.4.1.2 Ion trap (IT) instruments

Ion trap analyzers are able to "trap" ions in a physical device allowing them to build up over time (71) and are capable of performing high throughput peptide analysis (71, 77). IT analyzers include the three dimensional ion trap (QIT) and the linear ion trap (LIT) (77). QIT analyzers have few limitations including limited ion trapping capacity and low mass accuracy, whilst LIT enjoy higher ion trapping capacity (almost 10 times), higher sensitivity and high ion scanning and data acquisition speed however, mass accuracy and resolution are yet limited (70). In addition, LIT instruments can perform MS^n analyses that allow efficient characterization of peptide PTMs (71). Hybrid mass spectrometers containing LIT analyzers include Q-Q-LIT which possesses two quadrupoles and the LIT (70). In Q-Q-LIT, ion selection and fragmentation take place outside the ion trap followed by their entrapment in the ion trap for scanning and analysis (70). This spectrometer has precursor ion and neutral loss scanning abilities (71), and can be used in the identification of phosphopeptides (70). Neutral loss refers to the loss of a neutral fragment from a peptide ion that can take place during ion fragmentation (71). Neutral losses from phosphopeptides are discussed later in more details.

3.4.1.3 Fourier transform ion cyclotron resonance (FTICR) instruments

In FTICR, m/z ratios of analyte ions can be determined by measuring the frequency of ion motion under a static magnetic field (70). FTICR analyzers use the fast Fourier transform algorithm (79) for the generation of m/z spectra and are able to operate with high mass accuracy and resolution (70, 71). A hybrid mass spectrometer combining FTICR with LIT, is able to execute parallel MS and MS/MS measurements, generating high quality data that can be used for protein identification and quantitation. However, this hybrid mass spectrometer is limited by the slow rate of ion scanning (71, 77).

3.4.1.4 Orbitrap instruments

In orbitrap analyzers, ions are trapped by allowing them to orbit around a central electrode under static electric field (77). Another electrode is placed in the axial direction to the central electrode allowing the ions to oscillate in the axial direction and the m/z ratios of ions are determined by measuring the frequency of ion oscillation (77, 80). A hybrid mass spectrometer combining orbitrap with LIT (LTQ-Orbitrap) enjoys the high mass accuracy and high mass resolution of the orbitrap, and the fast operation of ion scanning and data acquisition of LIT (77). Furthermore, LTQ-Orbitrap is able to execute parallel MS and MS/MS measurements (77).

3.4.2 Phosphopeptide fragmentation

As mentioned above, tandem mass spectrometry involves the selective isolation of peptide ions (precursor ions) generated by ionizing techniques, fragmentation of precursor ions into product ions and the subsequent analysis of these product ions to generate fragmentation data from which structural information can be inferred, such as amino acid sequences and phosphorylation sites (13). More informative data can be generated by performing MS^n analysis where (n) cycles of product ion selection and fragmentation can be performed (13). The quality of the analytical data produced by MS/MS is greatly influenced by the chemical properties of the precursor ion and the method used for precursor ion fragmentation (13). When a peptide ion is fragmented at the bonds forming its backbone structure, sequence product ions are generated and the type of these ions depends on the type of the dissociated backbone bond, as well as the presence of the C- or N-terminal sides in the product ion (13). The types of fragmented ions (Figure 7) (13) have been previously described in the literature (81, 82), where upon dissociation of the C_α-$C_{carbonyl}$ bonds product ions of a-type (carrying the N-terminal segment) or x-type (carrying the C-terminal segment) are generated, whilst dissociation of the amide bonds $C_{carbonyl}$-N generates b-type (carrying the N-terminal segment) or y-type (carrying the C-terminal segment) product ions, and finally, dissociation of the N-C_α bonds leads to the generation of product ions of c-type (carrying the N-terminal segment) or z-type (carrying the C-terminal segment) (13).

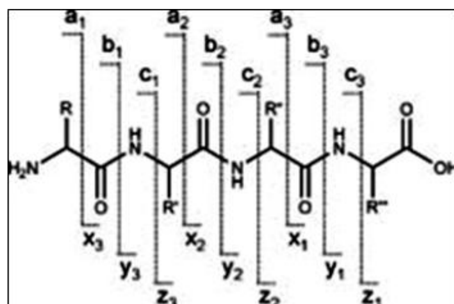

Figure 7 (13)**.** Types of sequence product ions generated by fragmentation of peptide backbone bonds during MS/MS analysis.

In addition to the dissociation of backbone bonds, dissociation of side chain bonds can take place resulting in the generation of nonsequence ions which do not provide sequence information, but can provide information regarding amino acid composition (13). For example, the dissociation of the phosphoester bond leads to the generation of nonsequence ions which can be used to indicate the presence of phosphorylation (13). Excessive generation of nonsequence ions can nevertheless, interfere with the identification of peptides and the determination of phosphorylation sites (13). There are many dissociation techniques that are employed in MS/MS strategies for phosphopeptide structural characterization that are discussed in details in this review (13). In this section however, the most established dissociation techniques in phosphoproteomic analysis are discussed which involve collision induced dissociation (CID) and the electron-based dissociation techniques: electron capture dissociation (ECD) and electron transfer dissociation (ETD) (13).

3.4.2.1 Collision-Induced dissociation (CID)

CID is the most utilized method for peptide ion fragmentation in MS/MS experiments, which are typically performed in positive ion mode where peptide ions gain protons and become positively charged (13, 15, 83). The underlying fragmentation mechanism of CID involves colliding precursor peptide ions with an inert gas, such as helium, argon or nitrogen which causes the precursor ions to vibrate by converting the translational energy of the coming precursor ion into internal vibrational energy (13, 15). When the vibrational energy of the precursor ion overcomes the activation energy of each bond they break causing ion fragmentation (13). Ideally, homologous population of fragmented ions from a given precursor ion are generated by the distribution of the acquired vibrational energy along the

entire backbone of the precursor peptide ion allowing the readout of such fragmented population ions to determine the primary amino acid sequence of the precursor peptide (4). The weaker the bond the lower the activation energy and the faster it breaks (13). This is of particular interest when comparing the CID-induced fragmentation of non-phosphorylated peptides to that of phosphorylated peptides. Protonation of the amide nitrogen weakens the amide bond and becomes the principal site of fragmentation generating product ions of b- and y-type (4, 15). However, in phosphopeptides, since the phosphoester bond is a weaker bond (with activation energy of < 20 Kcal/mol) than the amide bond (with activation energy of 40 kcal /mol), the phosphoester bond is preferentially broken imparting selective fragmentation of phosphopeptides (13). As a consequence, neutral mass loss from a phosphopeptide of 80 Da or 98 Da corresponding to the loss of HPO_3 and H_3PO_4 respectively takes place (15). In a CID spectrum, these neutral losses can dominate and overshadow the presence of other informative sequences which make it difficult to identify the amino acid sequences of peptides (15). Moreover, sequence product ions can undergo further fragmentation resulting in nonsequence neutral losses (13). Since neutral losses from phosphopeptides presents an obstacle in phosphopeptide characterization and localization of phosphorylation sites, understanding the factors that control the extent of neutral losses is therefore important. These factors include the precursor ion primary sequence and charge state, the type of the phosphorylated amino acid residue, the strength of collision energy generating product ions which can be sequence and nonsequence ions and the type of the mass spectrometer (15). The extent of neutral losses differs among phosphorylated serine, threonine and tyrosine amino acid residues, where neutral losses from a phosphorylated tyrosine is the least frequent with a typical neutral loss of HPO_3, followed by threonine with neutral losses of H_3PO_4 or HPO_3, while H_3PO_4 neutral loss is most frequent with phosphorylated serine (15). Complete protonation of the amide bond reduces the competition of the phosphoester bond directed fragmentation, which is influenced by the charge state of the precursor ion and the frequent presence of basic residues (arginine, histidine and lysine) in the amino acid sequence of the precursor ion (15). When CID is performed in collision cell mass spectrometers, the amount of collision energy is higher than that produced with resonant excitation CID performed in ion trap mass spectrometers which exhibit lower energy of excitation and therefore lower energy pathways, such as neutral losses are not preferred (15, 84, 85). Moreover, another limitation hampering the

correct identification of phosphorylation sites from product ions generated by resonance excitation CID, is the "phosphate group transfer reaction", where in a study performed by Palumbo and coworkers (86) on 33 potential phosphorylation sites-containing synthetic peptides that where characterized by CID-MS/MS and MS^3 in a quadrupole ion trap, 45% of these peptides generated product ions that experienced the transfer of a phosphate group from a protonated phosphorylated residue (tyrosine, threonine or serine) to a non-phosphorylated residue. Phosphate group transfer to amino acid residues with hydroxyl and carboxyl functional groups has also been demonstrated in another study with CID-MS/MS (87). The possibility of incorrect identifications of phosphorylation sites can be reduced by using high energy implementations such as with QqQ or qTOF mass spectrometers as to avoid the favorability of low energy transfer pathways or by using conditions allowing high charge state of the precursor ions when using ion trap mass spectrometers (86). Moreover, other dissociation methods can be used to avoid loss of the phosphate groups encountered with CID, such as electron capture dissociation (ECD) or electron transfer dissociation (ETD) (13) which are discussed below. In negative ion mode, CID-MS/MS precursor ions are deprotonated (83) and a characteristic PO_3^- (79 Da) ion is generated as a result of phosphoester bond cleavage of negatively charged peptide precursor ions containing phosphorylated serine, threonine and tyrosine residues (13, 83). Other charged losses corresponding to the loss of PO_2^- (63 Da) or $H_2PO_4^-$ (97 Da) can also take place (13). Because the interpretation of spectra generated by positive ion mode MS/MS is easier than those generated with negative ion mode MS/MS spectra, polarity switching during sample analysis or the performance of additional positive ion mode analysis can be achieved to generate positive ion mode spectra (13, 83).

3.4.2.2 Electron-based dissociation

Electron-based dissociation methods include electron capture dissociation (ECD) and electron transfer dissociation (ETD) which are able to fragment precursor peptide ions into product ions with intact post translational modifications (as phosphorylation and glycosylation) overcoming the neutral losses associated with CID (15). Both ECD and ETD are radical-triggered fragmentation methods, where in ECD a protonated multiply charged precursor peptide ion captures a low energy electron exothermically, resulting in the cleavage of the N-C_α bond generating c- and z-type product ions (13, 15). The reason that

ECD can maintain phosphate groups on the product ions is because dissociation takes place faster than the distribution of the vibrational energy along the backbone avoiding the favorable lower energy pathways as that of neutral losses (13) which have made it possible to clearly designate phosphorylation sites on phosphopeptides using ECD-MS/MS (88, 89). With ETD fragmented product ions of c- and z-type, similar to ECD, are generated via the dissociation of the N-C_α bond however, the dissociation mechanism involves the transfer of an electron from a radical anion with low electron affinity to the multiply protonated precursor peptide ion (13). Successful phosphopeptide identification with localization of phosphorylation sites have been achieved with ETD-based precursor ion fragmentation (90-92). Comparing the efficiency of phosphopeptide identification via ETD and CID, revealed that ETD is more efficient in identifying precursor ions with high charge densities, while CID is more efficient in the identification of phosphopeptides with low charge densities (90). However, studies suggest that comprehensive characterization of phosphopeptides can be achieved by the complementary use of ETD and CID (89, 90, 92).

3.5 Choice of instrumentation

When performing a phosphoproteomic study, the choice of the instrument is dictated by the application of interest (93). For example, when performing a large-scale phosphoproteomic study where thousands of phosphopeptides are to be identified with the localization of phosphorylation sites, the investigator wishes to perform parallel analysis of peptides in a rapid high throughput operation, and at the same time maintaining high mass accuracy and avoidance of phosphate neutral losses (15, 93). To circumvent the problem of averted peptide identification due to extensive neutral losses, as frequently seen with CID-induced fragmentation (13, 15), several strategies can be implemented. For example, MS^3 analysis can be performed with ion trap mass spectrometers where the isolation and fragmentation of a neutral loss fragment ion in a third MS stage (MS^3) can take place and the sequence of this fragment ion can be revealed as the dominance by the neutral loss peak is avoided, since the phosphate neutral loss already took place in the previous stage of ion isolation and fragmentation (15, 93). However, with multistage activation, the number of identified peptides can be reduced due to the implementation of extra activation steps (15). The use of ESI-based MS instruments instead of MALDI-based MS instruments can reduce phosphate neutral losses during fragmentation because of the predominant

generation of singly charged precursor ions with MALDI that may exhibit high rates of phosphate neutral losses during fragmentation due to low proton mobility, in contrast to ESI which predominantly generates multiply charged precursor ions (15). Substituting CID as the fragmentation technique with electron based dissociation techniques, such as ECD and ETD can prevent the generation of neutral-loss dominated spectra but they can be limited however, by the fact that they are more efficient with high charge density precursor ions (15). Moreover reducing the generation of phosphate neutral losses can be achieved with the use of Q-TOF instruments employing a high energy collision cell that favors the generation of sequence product ions as explained above (15, 93). Instrument choice can also be influenced by the quantification method (93). For instance, in SILAC experiments (stable isotope labeling of amino acids in culture), a quantitative approach of proteomic studies that will be discussed in the next chapter, the samples to be analyzed are highly complex because different samples are combined and analyzed in the same MS run (17). Therefore, such samples require MS instruments with high mass accuracy and high resolution such as hybrid mass spectrometers of FTICR with LIT or LTQ-Orbitrap (93).

Chapter 4: Quantitative phosphoproteomics

Quantifying changes in the phosphorylation state of a cell before and after specific cellular stimulations enables the investigation of (de)phosphorylation-mediated cellular responses toward specific stimuli (93). Quantifying the changes in cellular phosphorylation events can be performed using similar MS-based quantification methods employed in quantitative proteomics, which include stable isotope labeling quantitation, and label-free quantitation (17, 93). However, quantitative phosphoproteomics can be challenging because of several reasons (17). These include the necessity to differentiate true changes that occur in the phosphorylation state of a cell after a specific stimulus from changes that can occur as a result of changes in protein abundance which can be influenced by the cell type, variable conditions and the duration of investigation times (17). It is therefore important to normalize for protein expression to detect the true changes in phosphorylation states (17, 94). Moreover, the dynamic nature of cellular phosphorylation as a result of cellular translocation of phosphorylated proteins or changes in protein phosphorylated state regulated by kinases and phosphatases making quantitative phosphoproteomics additionally challenging (17). In this chapter MS-based methodologies for quantitative phosphoproteomics will be discussed with examples of some of the recent applications of each method.

4.1 Stable isotope labeling

Quantitative comparison of protein abundances between two cell populations can be achieved with stable isotope labeling (Figure 8.a) (95) which involves labeling the proteome of one cell population with a stable heavy isotope (e.g. ^{2}H, ^{13}C, ^{15}N or ^{18}O) and labeling the proteome of the other cell population with the corresponding natural isotope (5, 96). Peptides derived from both the heavy isotope-labeled and natural isotope-labeled cell populations are mixed and analyzed in the same MS run (5, 96). Each peptide will have two versions one labeled with the heavy isotope and one labeled with the light isotope where both are similar in their chemical and physical properties but are different in their masses producing two distinguishable peaks appear as doublets in a mass spectrum (5, 96). Comparing the relative intensities of each of the doublet peaks indicates the relative abundance or quantity of each peptide version in the two cellular populations compared (5).

This approach provide accurate quantification measurement because it allows simultaneous measurement of the two populations in the same run instead of different MS runs during which variations of protein behavior can be encountered (17, 96). Incorporation of stable isotopes can be achieved by metabolic labeling or chemical labeling (5).

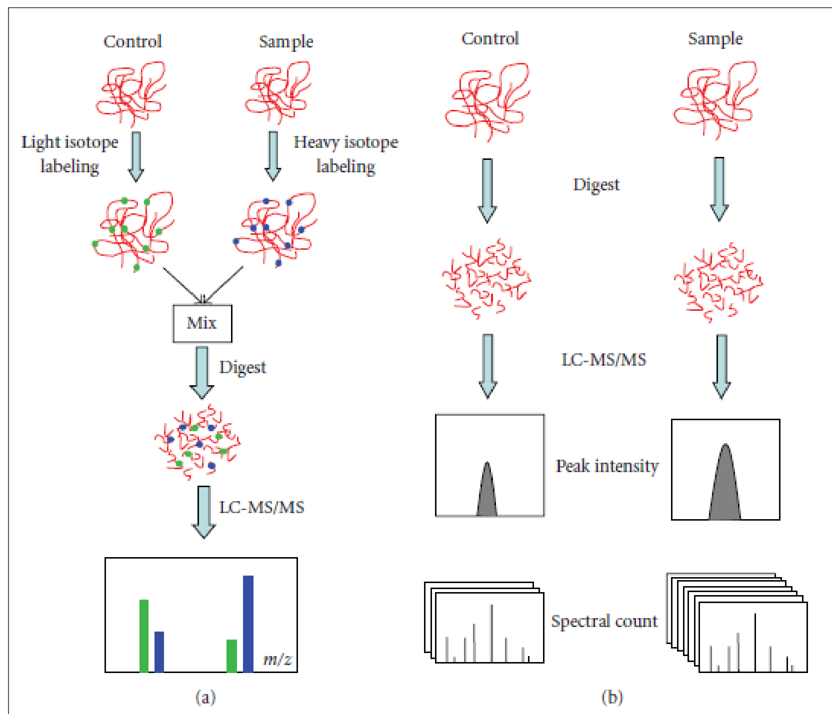

Figure 8 (95). Quantitative phosphoproteomics methodologies include: a) Stable isotope labeling quantitation and b) label-free quantitation.

4.1.1 Metabolic labeling

Metabolic labeling involves the incorporation of the heavy stable isotope label into proteins of growing cells (5). One common approach of metabolic labeling is SILAC (stable isotope labeling of amino acids in culture) (97) where isotopic labeled amino acids are added to the culture media of cultured cells and are incorporated into cellular proteins through de novo protein synthesis and protein degradation (96). SILAC allow the simultaneous comparison of samples set at 2 or 3 different conditions (17). SILAC was recently applied in a study with the aim of differentiating true changes in the phosphorylation state of a cell from those reflected by altered protein expression to account for accurate interpretation of phosphoproteomic data (94). They performed simultaneous SILAC/MS quantitation analysis of protein expression and phosphorylation state in mutant yeast cells with gene deletions encoding two proteins in the MAPK pathway (FUS3 or STE7) and the wild type versions (94).

From these experiments they were able to obtain quantitative ratios of 12,907 unique phosphopeptides upon quantitative comparison of cells with FUS3 deletion and the corresponding wild type cells, and 10588 unique phosphopeptides upon quantitative comparison of cells with STE7 deletion and the corresponding wild type cells (94). SILAC was also recently described for quantitative phosphoproteomics in an *in vivo* approach to study insulin signaling in mouse liver cells (98).

4.1.2 Chemical labeling

Incorporation of stable isotopes can be achieved by chemical tagging of peptides by isotope mass tags that can be chemically reacted with cysteine residues, amino groups or carboxylic groups (5). The most common reagents for chemical tagging are iTRAQ and TMT which are isotope-containing amine reactive molecules with MS/MS reporter groups allowing simultaneous identification and determination of peptide abundance at the tandem mass spectrometry level (5, 17). Upon ion fragmentation in MS/MS these tags dissociate into reporter ions and protein quantity can be inferred from the intensities of these reporter ions in the MS/MS spectra (5). Chemical labeling with TMT or iTRAQ allow simultaneous comparison of various samples set at different conditions (up to 8), thus reducing the burden of lengthy times of analyses (17, 95).

4.2 Label-free quantitation

In label-free quantitation approaches (Figure 8.b) different samples to be quantitatively compared are prepared and analyzed separately with MS (95). Protein quantitation is performed on each individual sample and can be achieved by either measuring the changes of ion peak intensities in the chromatogram, for example changes in peak areas or heights, or by counting identified peptides in the MS/MS spectra (95). Changes in protein quantity are determined by comparing the different quantitation results of the different samples (95). To achieve accurate label-free quantitation the reproducibility of the results must be maintained with the inclusion of different replicates of one sample for analysis (18). For more discussion of label-free quantitation the reader is directed for these reviews (95, 99).

The feasibility of the various quantitative approaches described above are dependent on several factors including sample type and complexity where SILAC is preferred in cell model systems while chemical labeling and label-free quantitation is preferred in human material

samples or animal models (17). Label-free approaches can provide quicker results with reduced sample complexity (95), it nevertheless introduce sample variability since samples to be compared are prepared and analyzed separately (17). This sample variability and therefore more accurate measurements can be achieved with the stable isotope labeling approaches (17).

Chapter 5: Data Analysis

Emergence of developmental advances in software technologies in the field of proteomics in general and phosphoproteomics as the branch topic of our interest will optimize the knowledge of a particular investigation of interest especially with the massive amount of information that can be generated with all developing strategies in sample preparation and instrumentation techniques described above (4, 100). In this section an overview of phosphoproteomic data processing with the current challenges will be addressed while more discussion can be found in these texts (4, 15, 101).

5.1 Database search

The first step in analyzing phosphopeptide fragmentation spectra is to correlate these spectra to the corresponding peptide amino acid sequences for peptide identification which can be performed through searching in sequence database (18). Specific computer programs can be employed for this sequence database search including Sequest (102), Mascot (103), OMSSA (104), X!Tandem (105) and InsPecT (106), while Sequest and Mascot have especially shown good performance with phosphopeptide identification (18). These algorithms can function by comparing the generated MS spectra to precursor ions or fragment ions with known calculated masses obtained from a database search after defining certain parameters and these peptides are assigned scores to identify the best matching or best scoring peptide among all the possible peptides present in the database (100, 103). Using high mass accuracy mass spectrometers assists the performance of database searching algorithms by decreasing the number of peptides that the algorithm consider for scoring (101). After the database search, the accuracy of peptide identification can be determined by identifying the false positives through the use of statistical filtration tools (18) such as PeptideProphet (107) with a typical cutoff value of 0.9 (18) or the use of target-decoy database searching where the false discovery rate (FDR) can be estimated (108) with typical cutoffs between 1% and 5% (18). Another filtration criterion that can be used to validate an identified matched as a true positive is mass accuracy (101). However, most of these programs are optimized for the analysis of peptide fragments generated by CID and therefore are not optimal for analyzing fragments generated by alternative fragmentation

techniques, such as electron-based dissociation techniques like ETD which, as described above, have proven to be efficient in the characterization of phosphorylated peptides (100). As also previously described, CID generates fragment ions of b- and y-type while ETD generates fragment ions of c- and z-type and therefore compatibility with ETD spectra requires the analysis of c- and z-type fragment ions (109). Moreover, the design of aforementioned algorithms is based on the analysis of peptides generated with trypsin, while other digestion proteases such as LysN have emerged for high throughput phosphopeptide analysis (1, 100). Recently, database search tools specific for ETD spectra have been developed, including Z-core (110) however, it was not superior to the older OMSSA algorithm (110). Another recent database search tool is MS-GFDB that was shown to outweigh Mascot performance in the analysis of ETD data sets and peptides generated with LysN (100).

5.2 Phosphosite localization

Searching for phosphorylation sites can be a complicated task because a given single peptide can contain frequent serine, threonine or tyrosine phosphorylation sites increasing the chance of having multiple phosphorylation sites (15). Using Mascot or Sequest algorithms, a score is given for all possible phosphorylation sites in a given peptide, and searching for the candidate site that acquired the highest score (15). The level of confidence for crediting a given site the highest score can be determined by calculating the difference between the highest score and the second highest score, so called delta score, such that the higher the delta score the higher the level of confidence for correct site assignment (15). Several PTM scoring tools exist (111-114) that can assist in the identification of phosphorylation sites using probability-based analysis (4, 15). Moreover, as also described above correct assignment of phosphorylation sites can be challenged by phosphate group rearrangement and transfer reaction (86).

5.3 Data storage and mining

Several websites have been established for investigators to share and search phosphoproteomic data such as phosphorylation sites and the responsible kinases (4). A list of these websites can be found in this review (4). Algorithms that assist searching and mining data have been established, such as 'motif-x' which is a motif extraction algorithm

that can be used to discover phosphorylation motifs from a data set (115). In addition, investigators can search for potential phosphorylation sites present in peptide sequences using prediction programs such as PhosphoMotif (116) and others (4), whilst a prediction approach that can annotate kinases for specific phosphorylations was also established namely NetworKIN (117).

5.4 Sulfonation

Phosphorylation adds a mass of 79.96633 Da to the modified protein (4) that is sometimes rounded to 80 Da (13) while sulfonation adds a mass of 79.95681 Da to the modified protein (4) which clearly announces another added challenge in the field of phosphoproteomics to not confuse and clearly differentiate protein phosphorylation from protein sulfonation (4). The continuous advancements in the field of phosphoproteomics with the development of high resolving power mass spectrometers and accurate data analysis tools can assist the clear distinction between the two modifications (4).

Chapter 6: Chemical biology tools to study protein phosphorylation

Chemical biology tools have emerged over the years to offer many contributions to the field of protein phosphorylation that greatly influenced the understanding of mechanisms governing the phosphorylation process (11). These tools enabled the investigation of cellular activities of many kinases, the impact of site specific phosphorylation on protein structure and function and monitoring the changes in time and cellular place of phosphorylation events (11). However, these tools are incapable of studying wide phosphoproteome events and are challenged by the dynamic nature of the phosphorylation process (11, 118). Although, as described previously, the state-of-art mass spectrometry methodologies have emerged to successfully perform global phosphoproteomic studies where thousands of phosphopeptides are identified with the localization of phosphorylation sites in a variety of complex biological samples, (4) developments in the chemical biology tools for studying protein phosphorylation didn't cease, where in fact in the last decade several applications of chemical biology tools have been published for studying phosphorylation in whole cell lysates and hold the promise for studying signaling pathways (118). These recent developments of the chemical biology tools will be discussed in this chapter and the reader is directed for these reviews for further discussions (11, 118).

6.1 Peptide biosensors

The design of peptide biosensors make use of the binding of peptides phosphorylated by the action of kinases to protein interaction domains that take place in cellular communication networks (10). Incorporation of fluorescent labels to either of the peptides via solid phase peptide synthesis or the protein interaction domains via labeling techniques creates the peptide biosensors in which the fluorescence of the fluorescent label can respond to the phosphorylation activity (10, 11). The response of fluorescent labels can be changed in emission wavelength and fluorescence quantum yields (10). Peptide biosensors can be employed to report phosphorylation in *in vitro* studies or can be used to monitor activity of kinases in living cells if the fluorescent molecule used is compatible with cellular studies (10, 11). Peptide biosensors can be categorized into environmentally sensitive

sensors, deep-quench biosensors, self-reporting biosensors and metal-chelation enhanced fluorescence biosensors (Figure 9) (11).

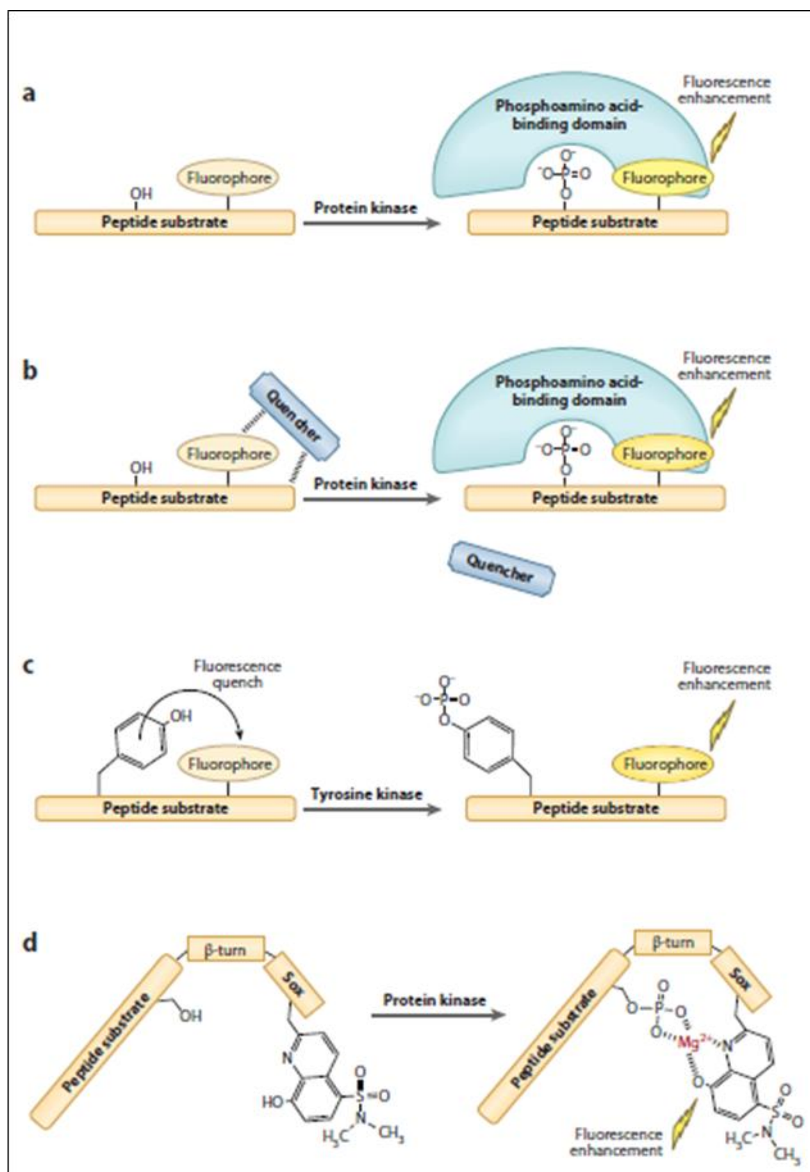

Figure 9 (11). Peptide biosensors include: a) Environmentally sensitive biosensor, b) Deep-quench biosensor, c) Self-reporting biosensor, d) Metal-chelation enhanced fluorescence biosensor (11).

6.1.1 Environmentally sensitive biosensors

In environmentally sensitive biosensors (Figure 9.a) (11) the fluorescence signal can respond to a change in solvent polarity present in the surrounding environment (10, 118). The change in the solvent polarity can be mediated by phosphorylation at a place close to the fluorescent label or can be due to binding of the phosphorylated peptide to a protein binding domain (118). An example of an environmental sensitive biosensor is a peptide

substrate of protein kinase C (PKC) containing the fluorophore reagent NBD that was discovered from a library prepared from various fluorescently labeled substrates of protein kinase C (PKC) (119), and was applied for monitoring intracellular kinase activity during cell division (120).

6.1.2 Deep-quench biosensors

Deep-quench biosensors (Figure 9.b) are based on kinase peptide substrates containing a fluorophore having its fluorescence suppressed or decreased by the presence of a quencher molecule that is non-covalently attached to the substrate (11). When the peptide carrying both the quencher and the fluorescent label is phosphorylated by the action of kinases the phosphorylated peptide binds a phosphoamino acid-specific binding domain separating the fluorescent molecule from the quencher causing an increase in fluorescence signal (11). An example of a deep-quench biosensor is a peptide substrate of protein kinase A (PKC) containing a pyrene fluorophore having its fluorescence quenched by a secondary dye that is non-covalently attached to the peptide (121). Upon incubating the peptide substrate containing the fluorophore and the quencher dye with PKA, ATP and a phosphoSer-binding domain, the fluorescence was increased by 64 folds (121).

6.1.3 Self-reporting biosensors

Self-reporting biosensors (Figure 9.c) are similar to the deep-quench biosensors but without the need for recruiting binding domains and that aromatic amino acid residues such as tyrosine can serve for quenching the fluorescence (11, 122). Phosphorylating the tyrosine residue disrupts the quenching of fluorescence resulting in an increase in the fluorescence signal (122).

6.1.4 Metal-chelation enhanced fluorescence biosensor

In metal-chelation enhanced fluorescence biosensors (Figure 9.d) the fluorescence signal can respond to physiological Mg^{2+} levels (11). The most established metal-chelation biosensors used are those based on the fluorophore sulfonamido-oxine (Sox) (11, 118) which is a non-natural amino acid fluorophore (Figure 10) (123) that upon Mg^{2+} chelation, an increase in its fluorescence signal take place (123). Many biosensors have been developed using Sox, which were generally designed to include an additional residue that

adjusts the metal binding, Ser/Thr residue and kinase recognition elements (11). The metal affinity is increased upon phosphorylation of the Ser/Thr residue causing generation of the fluorescence signal (11).

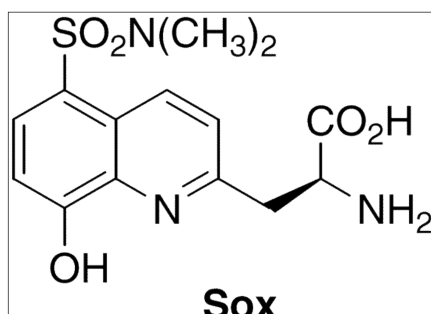

Figure 10 (123)**.** Structure of the fluorophore sulfonamido-oxine (Sox).

6.1.5 Genetically encoded protein sensors

Genetically encoded biosensors allow the continuous or real-time reporting of phosphorylation events inside cells (118). The most popular genetically encoded biosensors use changes in Forster resonance energy transfer (FRET) for monitoring phosphorylation activity (118). FRET involves the transfer of energy from an excited fluorophore to a neighboring acceptor fluorophore (124). The basic design of FRET-based biosensors include: two fluorescent proteins (donors and acceptors), a specific kinase substrate, and a phosphoamino acid binding domain (124). When the kinase phosphorylates its substrate, this causes the binding domain to bind the phosphorylated substrate which induces a conformational change that brings the two fluorophpores together allowing for the energy transfer (see Figure 11) (118, 124). In addition, FRET-based biosensors that allow monitoring of phosphatase enzymes have been developed as that used for monitoring the intracellular activity of Ca2+/calmodulin-dependent protein phosphatase (125).

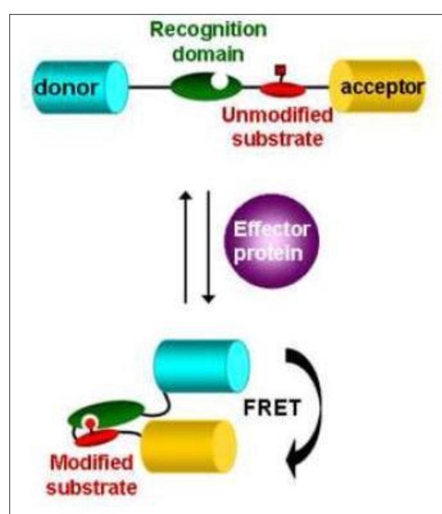

Figure 11 (124). Basic design of FRET (Forster resonance energy transfer)-based genetically encoded biosensors.

6.2 Small molecule mimics of phosphoamino acids

The reversal nature of protein phosphorylation presents an obstacle for *in vivo* studies of protein phosphorylation since proteins can undergo hydrolysis by the action of phosphatases (126). One approach to render proteins of interest resistance from phosphatase-mediated hydrolysis is to introduce a phosphoamino acid mimetic molecule into the sites that are normally phosphorylated in the protein (11, 118, 126). Small molecule mimics of phosphoamino acids include: non-hydrolyzable phosphoamino acids, caged phosphoamino acids and carboxyl-based non-hydrolyzable mimics (118). In caged phosphoamino acids the phosphate group is protected or "caged" by a photo-cleavable moiety that can be removed upon irradiation allowing the control over the time and place of phosphate group deprotection (127). Incorporation of *p*AA mimics at specific sites of proteins can be achieved by protein semi-synthesis or nonsense technology (118). In protein semi-synthesis the *p*AA mimic is first introduced to a peptide sequence using standard solid-phase peptide synthesis and the *p*AA-containing peptide is subsequently coupled to the protein of interest by chemical ligation (128). Nonsense technology makes use of a designed transfer RNA (tRNA) that is aminoacylated with the *p*AA mimic for its incorporation at specific sites of the protein of interest (129). An illustrative example for the use of *p*AAs in studying the effects of site-specific phosphorylation on the activity of cellular proteins is a study performed on one of the signal transducers and activators of transcription (STAT) proteins, STAT6 (130). They aimed to investigate the effect of deprotecting a photocaged phosphorylated tyrosine residue (pTyr641) on STAT6 dimerization and DNA binding in *in vitro* and its effect on STAT6 nuclear entrance in living cells (130). Recombinant STAT6 proteins were synthesized by protein semi-synthesis that contained phosphorylated Tyr641, non-phosphorylated Tyr 641 or photocaged Tyr641. In addition, hexa-histidine tags and fluorescence labels were attached for purification and visualization respectively. Deprotection of the photocaged Tyr641 allowed the investigators to control over STAT6 dimerization which is essential for DNA binding. Moreover, decaging of Tyr641 promoted STAT6 nuclear entrance which was visualized by the increase in the fluorescence signal in the nucleus accompanied by a decrease of the signal in the cytoplasm. This study highlights the impact of using a pAA mimic in the investigation of site-specific phosphorylation-mediated protein activity regulation.

6.3 Activity-based probes

Another chemical approach in studying the phosphorylation process is the activity-based protein profiling of kinases and phosphatases (118). The enzymatic activity of kinases and phosphatases present in complex protein populations can be monitored by activity-based probes (ABPs) (118). Activity-based probes (Figure 12) target and covalently modify the active sites of target enzymes, such as kinases or phosphatases, and are generally based on three main components comprising, a targeting device that renders the probe specificity toward a specific enzyme, a reactive molecule that covalently binds the active site of the enzyme and analytical tags or tools that allow the selective isolation of the target enzyme by immunoprecipitations or pull down assays and/or allow the visualization of the probe-labeled enzymes, for example fluorescent labels allow visualization by in-gel fluorescence scanning (118, 131). Several types of ABPs that are based on irreversible inhibitors of kinases or phosphatases were developed to target these enzymes (118), while one recent example of a probe targeting each are to be discussed. A peptide-based probe containing 2-fluoromethyl phosphotyrosine (2-FMPT) residue has been reported recently to target protein tyrosine phosphatases (PTPs) (132). 2-FMPT serves as a phosphotyrosine mimic molecule resembling the structure of a phosphotyrosine except for a 2- fluoromethyl group present in the aromatic ring allowing the probe recognition by PTPs (132). Moreover, higher specificity toward PTPs can be conquered through the inclusion of recognition elements present in a natural substrate of PTP (132). The mechanism of the probe binding is proposed to take place by binding the active site of a PTP followed by the generation of a reactive quinone methide intermediate that covalently alkylates the active site of the enzyme (132). ABPs based on either small molecule inhibitors or ATP analogues, the widely recognized co-substrate of kinases, can be used to target active site cysteine or lysine residues of kinases to monitor their activity (118, 133). An example of such probes is an adenosine derivative probe containing naphthalene-2,3-dialdehyde (NDA) where the adenosine moiety serves as the targeting device directing the probe to the ATP-binding site of the kinase, while the NDA acts as a cross-linking molecule to cross-link the target kinase with its substrate in cell lysates in an attempt to identify kinase-substrate pairs (133).

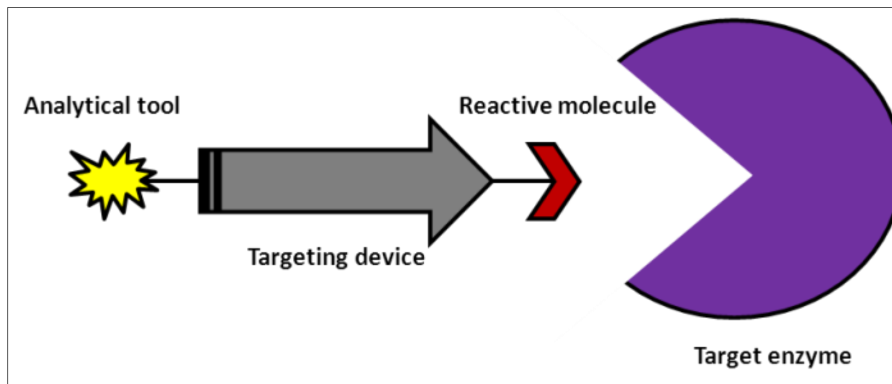

***Figure 12.* Schematic representation of activity-based probes.** Activity-based probes are mainly consisting of a targeting device that directs the probe toward the target enzyme, a reactive molecule that covalently binds the active site of the target enzyme and analytical tools that allow the purification and/or visualization of the probe-modified enzymes.

Chapter 7: Challenges and future directions

As discussed in this thesis, various developments in enrichment techniques and MS technologies have allowed researches to investigate phosphorylation in a variety of biological contexts, with the identification and quantitation of 15,000-20,000 unique phosphorylation sites (4). Nevertheless, current technologies still suffer from limitations and understanding these limitations with the current strengths, will aid in the development of more efficient methods and strategies. For example, the current enrichment techniques still lead to the identification of overlapping populations of phosphopeptides (4). Developments in the technologies, as well as complementary setups combining different enrichment techniques and prior separation by chromatographic techniques can greatly influence the coverage of phosphoproteomes. Because of the dynamic nature and the substoichiometric levels of phosphopeptides, developments in MS instrumentations aiming to enhance sensitivity and dynamic range of measurements will lead to the identification of the low abundant of phosphopeptides. Conventional LC-MS analyses are performed in positive ion modes which employ acidic conditions that result in the suppression of phosphopeptides ionization owing to the negative charge of the phosphate group (83). Negative ion mode MS can therefore, hold the promise to overcome the ionization suppression of phosphorylated peptides. However, the main limitation of negative ion modes is that the generated ion fragmentation spectra are difficult to interpret (13, 83). Some strategies can be employed to overcome this limitation including polarity switching during sample analysis or the performance of additional positive ion mode analysis to generate positive ion mode spectra (13, 83). Moreover, activation of peptide anions with electrons as using ECD and ETD can generate more informative sequence ions making the interpretation of spectra more straightforward (134). Another important challenge facing phosphoproteomics, is the interpretation of the readily available data that are massively increased with advanced technologies in sample preparation strategies and MS instrumentation. The ability to correlate these data to biological questions is another important area that can benefit from further developments in bioinformatics tools.

Acknowledgment

The writer of this literature thesis would like to thank her supervisor Dr. Dirk Rijkers for his guidance through the writing process and all the valuable discussions he kindly offered to complete this work. The writer would like to thank Dr. Shabaz Mohammed for accepting to review this literature thesis.

References

1. Gauci S, Helbig AO, Slijper M, Krijgsveld J, Heck AJ, Mohammed S. Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. Anal Chem. 2009 Jun 1;81(11):4493-501.

2. Cohen P. The role of protein phosphorylation in human health and disease. the sir hans krebs medal lecture. Eur J Biochem. 2001 Oct;268(19):5001-10.

3. Beltran L, Cutillas PR. Advances in phosphopeptide enrichment techniques for phosphoproteomics. Amino Acids. 2012 Sep;43(3):1009-24.

4. Grimsrud PA, Swaney DL, Wenger CD, Beauchene NA, Coon JJ. Phosphoproteomics for the masses. ACS Chem Biol. 2010 Jan 15;5(1):105-19.

5. Macek B, Mann M, Olsen JV. Global and site-specific quantitative phosphoproteomics: Principles and applications. Annu Rev Pharmacol Toxicol. 2009;49:199-221.

6. Ubersax JA, Ferrell JE,Jr. Mechanisms of specificity in protein phosphorylation. Nat Rev Mol Cell Biol. 2007 Jul;8(7):530-41.

7. Sacco F, Perfetto L, Castagnoli L, Cesareni G. The human phosphatase interactome: An intricate family portrait. FEBS Lett. 2012 Aug 14;586(17):2732-9.

8. Cutillas PR, Jorgensen C. Biological signalling activity measurements using mass spectrometry. Biochem J. 2011 Mar 1;434(2):189-99.

9. Sefton BM, Shenolikar S. Overview of protein phosphorylation. Curr Protoc Protein Sci. 2001 May;Chapter 13:Unit13.1.

10. Sharma V, Wang Q, Lawrence DS. Peptide-based fluorescent sensors of protein kinase activity: Design and applications. Biochim Biophys Acta. 2008 Jan;1784(1):94-9.

11. Tarrant MK, Cole PA. The chemical biology of protein phosphorylation. Annu Rev Biochem. 2009;78:797-825.

12. Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. Trends Biochem Sci. 2002 Oct;27(10):514-20.

13. Palumbo AM, Smith SA, Kalcic CL, Dantus M, Stemmer PM, Reid GE. Tandem mass spectrometry strategies for phosphoproteome analysis. Mass Spectrom Rev. 2011 Jul-Aug;30(4):600-25.

14. Marcantonio M, Trost M, Courcelles M, Desjardins M, Thibault P. Combined enzymatic and data mining approaches for comprehensive phosphoproteome analyses: Application to cell signaling events of interferon-gamma-stimulated macrophages. Mol Cell Proteomics. 2008 Apr;7(4):645-60.

15. Boersema PJ, Mohammed S, Heck AJ. Phosphopeptide fragmentation and analysis by mass spectrometry. J Mass Spectrom. 2009 Jun;44(6):861-78.

16. Thingholm TE, Jensen ON, Larsen MR. Analytical strategies for phosphoproteomics. Proteomics. 2009 Mar;9(6):1451-68.

17. Kanshin E, Michnick S, Thibault P. Sample preparation and analytical strategies for large-scale phosphoproteomics experiments. Semin Cell Dev Biol. 2012 Jun 6.

18. Bodenmiller B, Aebersold R. Quantitative analysis of protein phosphorylation on a system-wide scale by mass spectrometry-based proteomics. Methods Enzymol. 2010;470:317-34.

19. Thingholm TE, Larsen MR, Ingrell CR, Kassem M, Jensen ON. TiO(2)-based phosphoproteomic analysis of the plasma membrane and the effects of phosphatase inhibitor treatment. J Proteome Res. 2008 Aug;7(8):3304-13.

20. Ahmed FE. Sample preparation and fractionation for proteome analysis and cancer biomarker discovery by mass spectrometry. J Sep Sci. 2009 Mar;32(5-6):771-98.

21. Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol. 2004 Sep;5(9):699-711.

22. Raijmakers R, Neerincx P, Mohammed S, Heck AJ. Cleavage specificities of the brother and sister proteases lys-C and lys-N. Chem Commun (Camb). 2010 Dec 14;46(46):8827-9.

23. Taouatas N, Drugan MM, Heck AJ, Mohammed S. Straightforward ladder sequencing of peptides using a lys-N metalloendopeptidase. Nat Methods. 2008 May;5(5):405-7.

24. Taouatas N, Heck AJ, Mohammed S. Evaluation of metalloendopeptidase lys-N protease performance under different sample handling conditions. J Proteome Res. 2010 Aug 6;9(8):4282-8.

25. McNulty DE, Annan RS. Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. Mol Cell Proteomics. 2008 May;7(5):971-80.

26. Di Palma S, Hennrich ML, Heck AJR, Mohammed S. Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. Journal of Proteomics. 2012 7/16;75(13):3791-813.

27. Andersson L, Porath J. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem. 1986 4;154(1):250-4.

28. Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. Anal Chem. 2004 Jul 15;76(14):3935-43.

29. Zhang X, Ye J, Jensen ON, Roepstorff P. Highly efficient phosphopeptide enrichment by calcium phosphate precipitation combined with subsequent IMAC enrichment. Mol Cell Proteomics. 2007 Nov;6(11):2032-42.

30. Ruse CI, McClatchy DB, Lu B, Cociorva D, Motoyama A, Park SK, et al. Motif-specific sampling of phosphoproteomes. J Proteome Res. 2008 May;7(5):2140-50.

31. Posewitz MC, Tempst P. Immobilized gallium(III) affinity chromatography of phosphopeptides. Anal Chem. 1999 Jul 15;71(14):2883-92.

32. Aryal UK, Olson DJ, Ross AR. Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests. J Biomol Tech. 2008 Dec;19(5):296-310.

33. Kinoshita E, Yamada A, Takeda H, Kinoshita-Kikuta E, Koike T. Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins. J Sep Sci. 2005 Feb;28(2):155-62.

34. Feng S, Ye M, Zhou H, Jiang X, Jiang X, Zou H, et al. Immobilized zirconium ion affinity chromatography for specific enrichment of phosphopeptides in phosphoproteome analysis. Mol Cell Proteomics. 2007 Sep;6(9):1656-65.

35. Neville DC, Rozanas CR, Price EM, Gruis DB, Verkman AS, Townsend RR. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. Protein Sci. 1997 Nov;6(11):2436-45.

36. Anguenot R, Yelle S, Nguyen-Quoc B. Purification of tomato sucrose synthase phosphorylated isoforms by fe(III)-immobilized metal affinity chromatography. Arch Biochem Biophys. 1999 May 1;365(1):163-9.

37. Li S, Dass C. Iron(III)-immobilized metal ion affinity chromatography and mass spectrometry for the purification and characterization of synthetic phosphopeptides. Anal Biochem. 1999 May 15;270(1):9-14.

38. Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, et al. Phosphoproteome analysis by mass spectrometry and its application to saccharomyces cerevisiae. Nat Biotechnol. 2002 Mar;20(3):301-5.

39. Tsai CF, Wang YT, Chen YR, Lai CY, Lin PY, Pan KT, et al. Immobilized metal affinity chromatography revisited: PH/acid control toward high selectivity in phosphoproteomics. J Proteome Res. 2008 Sep;7(9):4058-69.

40. Salomon AR, Ficarro SB, Brill LM, Brinker A, Phung QT, Ericson C, et al. Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. Proc Natl Acad Sci U S A. 2003 Jan 21;100(2):443-8.

41. Lee J, Xu Y, Chen Y, Sprung R, Kim SC, Xie S, et al. Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. Mol Cell Proteomics. 2007 Apr;6(4):669-76.

42. Kim JE, Tannenbaum SR, White FM. Global phosphoproteome of HT-29 human colon adenocarcinoma cells. J Proteome Res. 2005 Jul-Aug;4(4):1339-46.

43. Stewart II, Thomson T, Figeys D. 18O labeling: A tool for proteomics. Rapid Commun Mass Spectrom. 2001;15(24):2456-65.

44. Thingholm TE, Jensen ON, Robinson PJ, Larsen MR. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. Mol Cell Proteomics. 2008 Apr;7(4):661-71.

45. Ye J, Zhang X, Young C, Zhao X, Hao Q, Cheng L, et al. Optimized IMAC-IMAC protocol for phosphopeptide recovery from complex biological samples. J Proteome Res. 2010 Jul 2;9(7):3561-73.

46. Cantin GT, Yi W, Lu B, Park SK, Xu T, Lee JD, et al. Combining protein-based IMAC, peptide-based IMAC, and MudPIT for efficient phosphoproteomic analysis. J Proteome Res. 2008 Mar;7(3):1346-51.

47. Choudhary C, Olsen JV, Brandts C, Cox J, Reddy PN, Bohmer FD, et al. Mislocalized activation of oncogenic RTKs switches downstream signaling outcomes. Mol Cell. 2009 Oct 23;36(2):326-39.

48. Aryal UK, Ross AR. Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. Rapid Commun Mass Spectrom. 2010 Jan;24(2):219-31.

49. Jensen SS, Larsen MR. Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. Rapid Commun Mass Spectrom. 2007;21(22):3635-45.

50. Rogers LD, Fang Y, Foster LJ. An integrated global strategy for cell lysis, fractionation, enrichment and mass spectrometric analysis of phosphorylated peptides. Mol Biosyst. 2010 May;6(5):822-9.

51. Ballif BA, Villen J, Beausoleil SA, Schwartz D, Gygi SP. Phosphoproteomic analysis of the developing mouse brain. Mol Cell Proteomics. 2004 Nov;3(11):1093-101.

52. Lim KB, Kassel DB. Phosphopeptides enrichment using on-line two-dimensional strong cation exchange followed by reversed-phase liquid chromatography/mass spectrometry. Anal Biochem. 2006 Jul 15;354(2):213-9.

53. Villen J, Gygi SP. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc. 2008;3(10):1630-8.

54. Lemeer S, Pinkse MW, Mohammed S, van Breukelen B, den Hertog J, Slijper M, et al. Online automated in vivo zebrafish phosphoproteomics: From large-scale analysis down to a single embryo. J Proteome Res. 2008 Apr;7(4):1555-64.

55. Han G, Ye M, Zhou H, Jiang X, Feng S, Jiang X, et al. Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography. Proteomics. 2008 Apr;8(7):1346-61.

56. Dai J, Wang LS, Wu YB, Sheng QH, Wu JR, Shieh CH, et al. Fully automatic separation and identification of phosphopeptides by continuous pH-gradient anion exchange online coupled with reversed-phase liquid chromatography mass spectrometry. J Proteome Res. 2009 Jan;8(1):133-41.

57. Hennrich ML, van den Toorn HW, Groenewold V, Heck AJ, Mohammed S. Ultra acidic strong cation exchange enabling the efficient enrichment of basic phosphopeptides. Anal Chem. 2012 Feb 21;84(4):1804-8.

58. Washburn MP, Wolters D, Yates JR,3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001 Mar;19(3):242-7.

59. Zhang G, Neubert TA. Use of detergents to increase selectivity of immunoprecipitation of tyrosine phosphorylated peptides prior to identification by MALDI quadrupole-TOF MS. Proteomics. 2006;6(2):571-8.

60. Kettenbach AN, Gerber SA. Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: Application to general and phosphotyrosine-specific phosphoproteomics experiments. Anal Chem. 2011 Oct 15;83(20):7635-44.

61. Pandey A, Podtelejnikov AV, Blagoev B, Bustelo XR, Mann M, Lodish HF. Analysis of receptor signaling pathways by mass spectrometry: Identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. Proc Natl Acad Sci U S A. 2000 Jan 4;97(1):179-84.

62. Steen H, Jebanathirajah JA, Rush J, Morrice N, Kirschner MW. Phosphorylation analysis by mass spectrometry: Myths, facts, and the consequences for qualitative and quantitative measurements. Mol Cell Proteomics. 2006 Jan;5(1):172-81.

63. Adamczyk M, Gebler JC, Wu J. Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry. Rapid Commun Mass Spectrom. 2001;15(16):1481-8.

64. Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol. 2001 Apr;19(4):379-82.

65. Poot AJ, Ruijter E, Nuijens T, Dirksen EH, Heck AJ, Slijper M, et al. Selective enrichment of ser-/Thr-phosphorylated peptides in the presence of ser-/Thr-glycosylated peptides. Proteomics. 2006 Dec;6(24):6394-9.

66. Li W, Backlund PS, Boykins RA, Wang G, Chen HC. Susceptibility of the hydroxyl groups in serine and threonine to beta-elimination/Michael addition under commonly used moderately high-temperature conditions. Anal Biochem. 2003 Dec 1;323(1):94-102.

67. Pinkse MW, Mohammed S, Gouw JW, van Breukelen B, Vos HR, Heck AJ. Highly robust, automated, and sensitive online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in drosophila melanogaster. J Proteome Res. 2008 Feb;7(2):687-97.

68. Nie S, Dai J, Ning ZB, Cao XJ, Sheng QH, Zeng R. Comprehensive profiling of phosphopeptides based on anion exchange followed by flow-through enrichment with titanium dioxide (AFET). J Proteome Res. 2010 Sep 3;9(9):4585-94.

69. Alcolea MP, Cutillas PR. In-depth analysis of protein phosphorylation by multidimensional ion exchange chromatography and mass spectrometry. Methods Mol Biol. 2010;658:111-26.

70. Yates JR,3rd. Mass spectral analysis in proteomics. Annu Rev Biophys Biomol Struct. 2004;33:297-316.

71. Domon B, Aebersold R. Mass spectrometry and protein analysis. Science. 2006 Apr 14;312(5771):212-7.

72. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 1989 Oct 6;246(4926):64-71.

73. Hillenkamp F, Karas M, Beavis RC, Chait BT. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. Anal Chem. 1991 Dec 15;63(24):1193A-203A.

74. Cech NB, Enke CG. Practical implications of some recent studies in electrospray ionization fundamentals. Mass Spectrom Rev. 2001 Nov-Dec;20(6):362-87.

75. Kuyama H, Sonomura K, Nishimura O. Sensitive detection of phosphopeptides by matrix-assisted laser desorption/ionization mass spectrometry: Use of alkylphosphonic acids as matrix additives. Rapid Commun Mass Spectrom. 2008 Apr;22(8):1109-16.

76. Dong M, Wu M, Wang F, Qin H, Han G, Dong J, et al. Coupling strong anion-exchange monolithic capillary with MALDI-TOF MS for sensitive detection of phosphopeptides in protein digest. Anal Chem. 2010 Apr 1;82(7):2907-15.

77. Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: Approaches, advances, and applications. Annu Rev Biomed Eng. 2009;11:49-79.

78. Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, et al. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. Anal Chem. 2000 Feb 1;72(3):552-8.

79. Senko MW, Canterbury JD, Guan S, Marshall AG. A high-performance modular data system for fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom. 1996;10(14):1839-44.

80. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. The orbitrap: A new mass spectrometer. Journal of Mass Spectrometry. 2005;40(4):430-43.

81. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom. 1984 Nov;11(11):601.

82. Biemann K. Contributions of mass spectrometry to peptide and protein structure. Biomed Environ Mass Spectrom. 1988 Oct;16(1-12):99-111.

83. Witze ES, Old WM, Resing KA, Ahn NG. Mapping protein post-translational modifications with mass spectrometry. Nat Methods. 2007 Oct;4(10):798-806.

84. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods. 2007 Sep;4(9):709-12.

85. Zhang Y, Ficarro SB, Li S, Marto JA. Optimized orbitrap HCD for quantitative analysis of phosphopeptides. J Am Soc Mass Spectrom. 2009 Aug;20(8):1425-34.

86. Palumbo AM, Reid GE. Evaluation of gas-phase rearrangement and competing fragmentation reactions on protein phosphorylation site assignment using collision induced dissociation-MS/MS and MS3. Anal Chem. 2008 Dec 15;80(24):9735-47.

87. Edelson-Averbukh M, Shevchenko A, Pipkorn R, Lehmann WD. Gas-phase intramolecular phosphate shift in phosphotyrosine-containing peptide monoanions. Anal Chem. 2009 Jun 1;81(11):4369-81.

88. Sweet SM, Mardakheh FK, Ryan KJ, Langton AJ, Heath JK, Cooper HJ. Targeted online liquid chromatography electron capture dissociation mass spectrometry for the localization of sites of in vivo phosphorylation in human Sprouty2. Anal Chem. 2008 Sep 1;80(17):6650-7.

89. Sweet SM, Bailey CM, Cunningham DL, Heath JK, Cooper HJ. Large scale localization of protein phosphorylation by use of electron capture dissociation mass spectrometry. Mol Cell Proteomics. 2009 May;8(5):904-12.

90. Swaney DL, McAlister GC, Coon JJ. Decision tree-driven tandem mass spectrometry for shotgun proteomics. Nat Methods. 2008 Nov;5(11):959-64.

91. Yang F, Wu S, Stenoien DL, Zhao R, Monroe ME, Gritsenko MA, et al. Combined pulsed-Q dissociation and electron transfer dissociation for identification and quantification of iTRAQ-labeled phosphopeptides. Anal Chem. 2009 May 15;81(10):4137-43.

92. Swaney DL, Wenger CD, Thomson JA, Coon JJ. Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. Proc Natl Acad Sci U S A. 2009 Jan 27;106(4):995-1000.

93. Nita-Lazar A, Saito-Benz H, White FM. Quantitative phosphoproteomics by mass spectrometry: Past, present, and future. Proteomics. 2008 Nov;8(21):4433-43.

94. Wu R, Dephoure N, Haas W, Huttlin EL, Zhai B, Sowa ME, et al. Correct interpretation of comprehensive phosphorylation dynamics requires normalization by protein expression changes. Mol Cell Proteomics. 2011 Aug;10(8):M111.009654.

95. Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol. 2010;2010:840518.

96. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc. 2006;1(6):2650-60.

97. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002 May;1(5):376-86.

98. Monetti M, Nagaraj N, Sharma K, Mann M. Large-scale phosphosite quantification in tissues by a spike-in SILAC method. Nat Methods. 2011 Jul 10;8(8):655-8.

99. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, et al. Less label, more free: Approaches in label-free quantitative mass spectrometry. Proteomics. 2011 Feb;11(4):535-53.

100. Kim S, Mischerikow N, Bandeira N, Navarro JD, Wich L, Mohammed S, et al. The generating function of CID, ETD, and CID/ETD pairs of tandem mass spectra: Applications to database search. Mol Cell Proteomics. 2010 Dec;9(12):2840-52.

101. Bakalarski CE, Haas W, Dephoure NE, Gygi SP. The effects of mass accuracy, data acquisition speed, and search algorithm choice on peptide identification rates in phosphoproteomics. Anal Bioanal Chem. 2007 Nov;389(5):1409-19.

102. Eng JK, McCormack AL, Yates III JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom. 1994 11;5(11):976-89.

103. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999 Dec;20(18):3551-67.

104. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, et al. Open mass spectrometry search algorithm. J Proteome Res. 2004 Sep-Oct;3(5):958-64.

105. Craig R, Beavis RC. TANDEM: Matching proteins with tandem mass spectra. Bioinformatics. 2004 Jun 12;20(9):1466-7.

106. Tanner S, Shu H, Frank A, Wang LC, Zandi E, Mumby M, et al. InsPecT: Identification of posttranslationally modified peptides from tandem mass spectra. Anal Chem. 2005 Jul 15;77(14):4626-39.

107. Keller A, Eng J, Zhang N, Li XJ, Aebersold R. A uniform proteomics MS/MS analysis platform utilizing open XML file formats. Mol Syst Biol. 2005;1:2005.0017.

108. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. 2007 Mar;4(3):207-14.

109. Coon JJ. Collisions or electrons? protein sequence analysis in the 21st century. Anal Chem. 2009 May 1;81(9):3208-15.

110. Sadygov RG, Good DM, Swaney DL, Coon JJ. A new probabilistic database search algorithm for ETD spectra. J Proteome Res. 2009 Jun;8(6):3198-205.

111. Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol. 2006 Oct;24(10):1285-92.

112. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell. 2006 Nov 3;127(3):635-48.

113. Ruttenberg BE, Pisitkun T, Knepper MA, Hoffert JD. PhosphoScore: An open-source phosphorylation site assignment tool for MSn data. J Proteome Res. 2008 Jul;7(7):3054-9.

114. Bailey CM, Sweet SM, Cunningham DL, Zeller M, Heath JK, Cooper HJ. SLoMo: Automated site localization of modifications from ETD/ECD mass spectra. J Proteome Res. 2009 Apr;8(4):1965-71.

115. Schwartz D, Gygi SP. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat Biotechnol. 2005 Nov;23(11):1391-8.

116. Amanchy R, Periaswamy B, Mathivanan S, Reddy R, Tattikota SG, Pandey A. A curated compendium of phosphorylation motifs. Nat Biotechnol. 2007 Mar;25(3):285-6.

117. Linding R, Jensen LJ, Ostheimer GJ, van Vugt MA, Jorgensen C, Miron IM, et al. Systematic discovery of in vivo phosphorylation networks. Cell. 2007 Jun 29;129(7):1415-26.

118. H S Lu C, Liu K, Tan LP, Yao SQ. Current chemical biology tools for studying protein phosphorylation and dephosphorylation. Chemistry. 2012 Jan 2;18(1):28-39.

119. Yeh RH, Yan X, Cammer M, Bresnick AR, Lawrence DS. Real time visualization of protein kinase activity in living cells. J Biol Chem. 2002 Mar 29;277(13):11527-32.

120. Dai Z, Dulyaninova NG, Kumar S, Bresnick AR, Lawrence DS. Visual snapshots of intracellular kinase activity at the onset of mitosis. Chem Biol. 2007 Nov;14(11):1254-60.

121. Sharma V, Agnes RS, Lawrence DS. Deep quench: An expanded dynamic range for protein kinase sensors. J Am Chem Soc. 2007 Mar 14;129(10):2742-3.

122. Wang Q, Cahill SM, Blumenstein M, Lawrence DS. Self-reporting fluorescent substrates of protein tyrosine kinases. J Am Chem Soc. 2006 Feb 15;128(6):1808-9.

123. Shults MD, Pearce DA, Imperiali B. Modular and tunable chemosensor scaffold for divalent zinc. J Am Chem Soc. 2003 Sep 3;125(35):10591-7.

124. Aye-Han NN, Ni Q, Zhang J. Fluorescent biosensors for real-time tracking of post-translational modification dynamics. Curr Opin Chem Biol. 2009 Oct;13(4):392-7.

125. Newman RH, Zhang J. Visualization of phosphatase activity in living cells with a FRET-based calcineurin activity sensor. Mol Biosyst. 2008 Jun;4(6):496-501.

126. Szewczuk LM, Tarrant MK, Cole PA. Protein phosphorylation by semisynthesis: From paper to practice. Methods Enzymol. 2009;462:1-24.

127. Mayer G, Heckel A. Biologically active molecules with a "light switch". Angew Chem Int Ed Engl. 2006 Jul 24;45(30):4900-21.

128. Muir TW. Semisynthesis of proteins by expressed protein ligation. Annu Rev Biochem. 2003;72:249-89.

129. Rodriguez EA, Lester HA, Dougherty DA. Improved amber and opal suppressor tRNAs for incorporation of unnatural amino acids in vivo. part 1: Minimizing misacylation. RNA. 2007 Oct;13(10):1703-14.

130. Lahiri S, Seidel R, Engelhard M, Becker CF. Photocontrol of STAT6 dimerization and translocation. Mol Biosyst. 2010 Dec;6(12):2423-9.

131. Ovaa H. Active-site directed probes to report enzymatic action in the ubiquitin proteasome system. Nat Rev Cancer. 2007 Aug;7(8):613-20.

132. Kalesh KA, Tan LP, Lu K, Gao L, Wang J, Yao SQ. Peptide-based activity-based probes (ABPs) for target-specific profiling of protein tyrosine phosphatases (PTPs). Chem Commun (Camb). 2010 Jan 28;46(4):589-91.

133. Liu K, Kalesh KA, Bing Ong L, Yao SQ. An improved mechanism-based cross-linker for multiplexed kinase detection and inhibition in a complex proteome. ChemBioChem. 2008;9(12):1883-8.

134. Yoo HJ, Wang N, Zhuang S, Song H, Hakansson K. Negative-ion electron capture dissociation: Radical-driven fragmentation of charge-increased gaseous peptide anions. J Am Chem Soc. 2011 Oct 26;133(42):16790-3.